

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2004, 17:40:11 ; Search time 11860.7 Seconds
(without alignments)
19051.434 Million cell updates/sec

Title: US-10-786-065-3_COPY_20000_26200

Perfect score: 6201
Sequence: 1 caccatggctcaggaacatc.....aagagtaacgaagaatgrrg 6201

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 32822875 seqs, 1821986598 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_esc1: *
2: gb_esc2: *
3: gb_hic: *
4: gb_esc3: *
5: gb_esc4: *
6: gb_esc5: *
7: gb_esc6: *
8: gb_gsa1: *
9: gb_gsa2: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	550.2	8.9	653	9 AG040253	AG040253 Pan trogl
2	413.6	6.7	710	4 BI091930	BI091930 602856713
3	412.8	6.7	517	6 CD521178	CD521178 AGENCOURT
4	412.8	6.7	583	5 BP871658	BP871658 BP871658
5	412.8	6.7	774	5 BU933084	BU933084 AGENCOURT
6	412.8	6.7	948	5 BU955838	BU955838 AGENCOURT
7	412	6.6	676	4 BI258333	BI258333 602968259
8	411.8	6.6	851	5 BU95722	BU95722 AGENCOURT
9	411	6.6	727	2 BE393340	BE393340 601311002
10	410.8	6.6	708	4 BI489733	BI489733 603032230
11	403.2	6.5	894	5 BU507729	BU507729 AGENCOURT
12	403	6.5	1180	3 BC005215	BC005215 Homo sapi
13	402.6	6.5	489	7 CN346858	CN346858 170005328
14	402.6	6.5	500	4 AI376271	AI376271 tae5h12.x
15	402.6	6.5	491	4 BM739173	BM739173 K-EST0008
16	402.6	6.5	599	1 AI418931	AI418931 tES2f08.x
17	402.6	6.5	601	4 BM753449	BM753449 K-EST0030
18	402.6	6.5	635	6 CB054131	CB054131 NISC_gm02
19	402.6	6.5	639	1 AI597626	AI597626 tni5g07.x
20	402.6	6.5	673	7 CN346851	CN346851 170006001
21	402.6	6.5	675	7 CN346862	CN346862 170006000
22	402.6	6.5	690	7 CN346845	CN346845 170005328
23	402.6	6.5	693	4 BG700992	BG700992 602682170
24	402.6	6.5	702	7 CN346850	CN346850 170005831

25	402.6	6.5	710	7 CN346828	CN346828 170005325
26	402.6	6.5	720	6 CF132267	CF132267 UI-HR-F00
27	402.6	6.5	721	4 BF984014	BF984014 602306941
28	402.6	6.5	721	6 CA391711	CA391711 c618D04.Y
29	402.6	6.5	725	4 BM762609	BM762609 K-EST0043
30	402.6	6.5	737	7 CN346865	CN346865 170006000
31	402.6	6.5	740	4 BI222628	BI222628 602940118
32	402.6	6.5	754	7 CN346857	CN346857 170005999
33	402.6	6.5	770	7 CN346837	CN346837 170005977
34	402.6	6.5	795	5 BQ221371	BQ221371 AGENCOURT
35	402.6	6.5	812	4 BG819344	BG819344 602781541
36	402.6	6.5	814	1 AU117598	AU117598 AU117598
37	402.6	6.5	822	1 AU110030	AU110030 DKE2686P
38	402.6	6.5	828	1 AU117451	AU117451 AU117451
39	402.6	6.5	849	2 BE747955	BE747955 601580366
40	402.6	6.5	844	1 AU124576	AU124576 AU124576
41	402.6	6.5	845	4 BI909653	BI909653 603066683
42	402.6	6.5	854	5 BU153052	BU153052 AGENCOURT
43	402.6	6.5	857	5 BU156219	BU156219 AGENCOURT
44	402.6	6.5	869	5 BQ232770	BQ232770 AGENCOURT
45	402.6	6.5	875	1 AU124769	AU124769 AU124769

ALIGNMENTS

RESULT 1
LOCUS AG040253 653 bp DNA linear GSS 01-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-017017.F, genomic survey sequence.
ACCESSION AG040253
VERSION AG040253.1 GI:16568978
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of library PTB
2 (bases 1 to 653)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22, Suenhiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpanzee@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB. This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKSI45
R Site 1 : SacI
R Site 2 : SacI
Location/Qualifiers
1.653
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-017017.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match 8.9%; Score 550.2; DB 9; Length 653;
Best Local Similarity 98.4%; Pred. No. 2.6e-37;

```
Matches 566; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 3144 AAATCCACACCTCAATTAATCTTCCAGCTCGGCTCCCAAGATGCTGGAATTACAGGT 3203
Db 647 AACCTCCACGCTCCATTAATCTTCCAGCTCGGCTCCCAAGATGCTGGAATTACAGGT 588
QY 3204 GTAGACCAAGGGTTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3262
Db 587 GTAGACCAAGGGTTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 528
QY 3263 GACAAGTACCGGGAACCTGACTCAACAGCTGCGGTTTAAGCTGAGCTGCTGCAATCT 3322
Db 527 GACAAGTACCGGGAACCTGACTCAACAGCTGCGGTTTAAGCTGAGCTGCTGCAATCT 468
QY 3323 AACATGATCAACCTGATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3382
Db 467 AACATGATCAACCTGATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 408
QY 3383 TCTAGTTAGCAATTAAGTAATCAAGTGTGATTAATGATGATGATGATGATGATGATG 3442
Db 407 TCTAGTTAGCAATTAAGTAATCAAGTGTGATTAATGATGATGATGATGATGATGATG 348
QY 3443 AATGTGAAATCACTAACAAGAAAGATTCACCTTTGTTGTTGTTGTTGTTGTTGTTGTT 3502
Db 347 AATGTGAAATCACTAACAAGAAAGATTCACCTTTGTTGTTGTTGTTGTTGTTGTTGTT 288
QY 3503 AAGGATGGGGCCAACTAACTGCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3562
Db 287 AAGGATGGGGCCAACTAACTGCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 228
QY 3563 TTAACAGAAATAGAGAGAAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3622
Db 227 TTAACAGAAATAGAGAGAAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 168
QY 3623 TTGAGGTCCAGGTGAGACAGATCTTCAAGGAAATGACAGTACGCACTTAAGAG 3682
Db 167 TTGAGGTCCAGGTGAGACAGATCTTCAAGGAAATGACAGTACGCACTTAAGAG 108
QY 3683 AATCTGTGCTCAGAAAGGAGCTGTGACTCGACG 3717
Db 107 AATCTGTGCTCAGAAAGGAGCTGTGACTCGACG 73

RESULT 2
LOCUS B1091930 710 bp mRNA linear EST 20-JUN-2001
DEFINITION 602856713F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4998218 5',
mRNA sequence.
ACCESSION B1091930
VERSION B1091930.1 GI:14510260
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 710)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLNI, at:
http://image.jnl.gov
Plate: L14M1026 row: 1 column: 03
High quality sequence stop: 707.
Location/Qualifiers
1..710
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4998218"
/cell_line="MGC36"
/lab_host="DH10B"
/clone_lib="NIH_MGC_10"
/note="Organ: cervix; Vector: pCMV-Sport6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5 kb. Library prepared by Life
Technologies."

ORIGIN
Query Match 6.7%; Score 413.6; DB 4; Length 710;
Best Local Similarity 92.7%; Pred. No. 5.8e-26;
Matches 434; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 16 CAATTTATGAAGCCAAATATACAGCTTAATAGTAATGAGTAAATCCAGAG 75
Db 178 CAATTTATGAAGCCAAATATACAGCTTAATAGTAATGAGTAAATCCAGAG 237
QY 76 AACTCCCTTTGTAAGATTTGTACAAAATTAATATGATGATGATGATGATGATGATG 135
Db 238 CACCCCTTTGTAAGATTTGTACAAAATTAATATGATGATGATGATGATGATGATG 297
QY 136 GAATGTGAACCCAAAGACCATATCAGGCTTGCAGAAATGGCAGATTCATATCATCA 195
Db 298 GAGTGTGACCCCAAGAGCATATCAGTGTACGCAAAATGGCAGATTCATATCATCA 357
QY 196 AAGTTATCCCTTCAAGAGCTTGCAGCCCTATATGATGCTTAAGAAATGAAAGCCCTC 255
Db 358 AAGTTATCCCTTCAAGAGCTTGCAGCCCTATATGATGCTTAAGAAATGAAAGCCCTC 417
QY 256 AGCCATCTGAAGAGACAGTGTACAGCAATTGATCAAAAAAGAAACCAAGGCCCTTCC 315
Db 418 ACCCCCGGAGAGACAGTGTACAGCAATTGATCAAAAAAGAAACCAAGGCCCTTCC 477
QY 316 CTTCCCCCAATCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 375
Db 478 CTTCCCCCAATCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 537
QY 376 AGCTGTAGAGCTCAAGTATCTGAAAGAAAGAGCTCCCATTCAGAGAAATTTATCTTAAG 435
Db 538 GTCTGTGAGACCTCAAGTATCTGAAAGAAAGAGCTCCCATTCAGAGAAATTTATCTTAAG 597
QY 436 ATACTGTAAATGATATTAATTTTGTACATTTGGAATATATATATGTTT 483
Db 598 ATACTGTAAATGATATTAATTTTGTACATTTGGAATATATATATGTTT 645

RESULT 3
LOCUS CD522178 517 bp mRNA linear EST 06-JUN-2003
DEFINITION AGENCOURT 14355240 NIH MGC 191 Homo sapiens cDNA clone
IMAGE:30410716 5', mRNA sequence.
ACCESSION CD522178
VERSION CD522178.1 GI:31453896
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 517)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Institutes / NIH
National Cancer Institute
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Narayan Bhat
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
```

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: NDCM201 row: F column: 05
 High quality sequence stop: 503.
 Location/Qualifiers

FEATURES

source

1..517
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30410716"
 /issue_type="Pooled"
 /lab_host="DH10B (TI phage-resistant)"
 /clone_lib="NIH_MGC_191"
 /note="Vector: pDNR-Lib; Site 1: SfiI (ggccatcggcc);
 Site 2: SfiI (ggccgcgcgcgc); Library is oligo-dT primed
 and directionally cloned. PBMC - Peripheral Blood
 Mononuclear Cells. RNA was pooled from 3/6hour stimulation
 with PMA and Ionomycin. 5' and 3' adaptors were used in
 cloning as follows: 5' adaptor sequence:
 5'-CACGCCCATTTATGCG-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCGCGCGCGCATG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.69
 kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 6.7%; Score 412.8; DB 6; Length 517;
 Best Local Similarity 92.1%; Pred. No. 8.3e-26;

Matches 446; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

1 CACCATGGGCTCAGCAATTTATGAAGCCAAATATACAGCCTTAATAATGATGTGAC 60
 23 CACGACGGCTCGACAAATTTATGAAGCCAAATATACAGCCTTAATAATGATGTGAC 82
 61 CTAAATATCCAGAGAACTCCCTTTGTAAGATTTGTAACAAAATTAATATGATGAG 120
 83 CTAAATATCCAGAGAACCCCTTTGTAAGATTTGTAACAAAATTAATATGATGAG 142
 121 TTAATAGTTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
 143 TTAATAGTTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 202
 181 ATTCAATATCATCAAGTTATCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAG 240
 203 ATTCAATATCATCAAGTTATCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAG 262
 241 ATGTGAAGAGCCCTCAGCATCTGAAGAGCAGTGTATACAGCAATGATCAAAAAGAAA 300
 263 ATATGAAGAGCCCTCAGCATCTGAAGAGCAGTGTATACAGCAATGATCAAAAAGAAA 322
 301 CCACAGGCGCTTCCCTTCCCATCTGATGATGATGATGATGATGATGATGATGATG 360
 323 CCACAGGCGCTTCCCTTCCCATCTGATGATGATGATGATGATGATGATGATGATG 382
 361 AAATTTTCTAGATACAGCTTGTAGAGCTCAAGTATCAAGTATCAAGTATCAAGTAT 420
 383 AAATTTTCTAGATACAGCTTGTAGAGCTCAAGTATCAAGTATCAAGTATCAAGTAT 442
 421 GAAATTTATCTAAGTATCTAAGTATCTAAGTATCTAAGTATCTAAGTATCTAAGTAT 479
 443 GAAATTTATCTAAGTATCTAAGTATCTAAGTATCTAAGTATCTAAGTATCTAAGTAT 502
 480 TTGT 483
 503 TTGT 506

RESULT 4
 BP871658

LOCUS BP871658 583 bp mRNA linear EST 27-JUL-2004
 DEFINITION BP871658 Sugano cDNA library, embryonal kidney Homo sapiens cDNA
 clone HRR01130, mRNA sequence.
 ACCESSION BP871658
 VERSION BP871658.1 GI:50716151
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Suzuki Y., Yamashita R., Shiota M., Sakakibara Y., Chiba J.,
 1 (bases 1 to 583)
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

TITLE
 JOURNAL
 COMMENT Database of Transcriptional Start Sites (DBTSS) for Comparative
 Studies of the Promoters of Human and Mouse Genes
 Unpublished (2004)
 Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES

source

1..583
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HRR01130"
 /issue_type="embryonal kidney"
 /cell_line="293"
 /clone_lib="Sugano cDNA library, embryonal kidney"

ORIGIN

Query Match 6.7%; Score 412.8; DB 5; Length 583;
 Best Local Similarity 92.1%; Pred. No. 7.7e-26;

Matches 446; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

1 CACCATGGGCTCAGCAATTTATGAAGCCAAATATACAGCCTTAATAATGATGTGAC 60
 77 CACGACGGCTCGACAAATTTATGAAGCCAAATATACAGCCTTAATAATGATGTGAC 136
 61 CTAAATATCCAGAGAACTCCCTTTGTAAGATTTGTAACAAAATTAATATGATGAG 120
 137 CTAAATATCCAGAGAACCCCTTTGTAAGATTTGTAACAAAATTAATATGATGAG 196
 121 TTAATAGTTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
 197 TTAATAGTTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 256
 181 ATTCAATATCATCAAGTTATCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAG 240
 257 ATTCAATATCATCAAGTTATCTTCAAGAGCTTCAAGAGCTTCAAGAGCTTCAAGAG 316
 241 ATGTGAAGAGCCCTCAGCATCTGAAGAGCAGTGTATACAGCAATGATCAAAAAGAAA 300
 317 ATATGAAGAGCCCTCAGCATCTGAAGAGCAGTGTATACAGCAATGATCAAAAAGAAA 376
 301 CCACAGGCGCTTCCCTTCCCATCTGATGATGATGATGATGATGATGATGATGATG 360
 377 CCACAGGCGCTTCCCTTCCCATCTGATGATGATGATGATGATGATGATGATGATG 436
 361 AAATTTTCTAGATACAGCTTGTAGAGCTCAAGTATCAAGTATCAAGTATCAAGTAT 420
 437 AAATTTTCTAGATACAGCTTGTAGAGCTCAAGTATCAAGTATCAAGTATCAAGTAT 496
 421 GAAATTTATCTAAGTATCTAAGTATCTAAGTATCTAAGTATCTAAGTATCTAAGTAT 479
 497 GAAATTTATCTAAGTATCTAAGTATCTAAGTATCTAAGTATCTAAGTATCTAAGTAT 556
 480 TTGT 483
 557 TTGT 560

RESULT 5
BU933084
LOCUS
DEFINITION
AGENCOURT_10473040 NIH_MGC_127 Homo sapiens cDNA clone
IMAGE:6673789 5', mRNA sequence.
BU933084
ACCESSION
BU933084.1 GI:24121903
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 774)
REFERENCE
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLCM2956 row: h column: 13
High quality sequence stop: 549.
Location/Qualifiers

FEATURES

source

1..774
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6673789"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_1ib="NIH_MGC_127"
/note="Vector: pDNR-Lib; Site 1: SfiI (ggccatcatggcc);
Site 2: SfiI (ggccgcctggcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell -
5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%,
salivary gland - 1.3%, and skin - 2.3%). 5' and 3'
adaptors were used in cloning as follows:
5'-AGCAGTGGATCAACGACAGCGCATGACGCGCG-3' and
5'-ATTCTAGAGCCCGACGCGCATG-CT(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 1-2 kb
size fraction (other fractions present in NIH_MGC_126 and
NIH_MGC_128). Library created in the laboratory of T.
Udell, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC
Library."

ORIGIN

Query Match 6.7%; Score 412.8; DB 5; Length 774;
Best Local Similarity 92.1%; Pred. No. 6.4e-26;
Matches 446; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

ORIGIN
Query Match 6.7%; Score 412.8; DB 5; Length 774;
Best Local Similarity 92.1%; Pred. No. 6.4e-26;
Matches 446; Conservative 0; Mismatches 37; Indels 1; Gaps 1;
1 CACCATGGCTCAGGCAATTTATGAAAGCCAAATATACAGCTTAAATAGATGTGAC 60
2 |||||
3 CACCATGGCTCAGGCAATTTATGAAAGCCAAATATACAGCTTAAATAGATGTGAC 89
4 |||||
5 CTAATATACCCAGAAAGAACCCCTTTGTAGATTTGTACAAAATTAATATAGATGAG 120
6 |||||
7 CTAATATACCCAGAAAGAACCCCTTTGTAGATTTGTACAAAATTAATATAGATGAG 149
8 |||||
9 CTAATATACCCAGAAAGAACCCCTTTGTAGATTTGTACAAAATTAATATAGATGAG 180
10 |||||
11 TTAATAGTTCTAATGGAATGTGAACCAAGAGCCATATCAGCGTAGCAAAATGGCAGA 180
12 |||||
13 TAAATAGTTCTAATGGAATGTGAACCAAGAGCCATATCAGCGTAGCAAAATGGCAGA 209
14 |||||
15 ATTCTATATATCAAAAGTTATCTTCAAGAGCTTCAGGCTTAATGATGTCAAAAGAA 240
16 |||||

DB 210 ATTCTATATATCAAAAGTTATCTTCAAGAGCTTCAGGCTTAATGATGTCTAAAGAA 269
OR 241 ATGTGAAGAGCCCTCAGCCATCTGAAGAGCAGTGTATACAGCAATTCATCAAAAGAAAA 300
270 ATATGAAGTCCCTCAGCCGCGGCGAAGAGCAGTGTATACAGCAATTCATCAAAAGAAAA 329
OR 301 CCACAGGCGCTTCCCTTCCCGCATATCTGATGAAGCAGTCTTCAATTTCCATGTAGT 360
330 CCACAGGCGCTTCCCTTCCCGCATATCTGATGAAGCAGTCTTCAATTTCCATGTAGT 389
OR 361 AAATTTTCTAGATAGCTGTAGAGCTCAAGTACTGGAAGAAAGCTCCCATTCAGAG 420
390 AAATTTTCTAGATAGCTGTAGAGCTCAAGTACTGGAAGAAAGCTCCCATTCAGAG 449
DB 421 GAAATTTATCTTAAATAGTCTGTAATGATGATCTTA-TTTTGTACATTTGGAATATATAG 479
450 GAAATTTATCTTAAATAGTCTGTAATGATGATCTTA-TTTTGTACATTTGGAATATATAG 509
OR 480 TTGT 483
510 TTGT 513

RESULT 6
BU955838
LOCUS
DEFINITION
AGENCOURT_10621087 NIH_MGC_107 Homo sapiens cDNA clone
IMAGE:6729290 5', mRNA sequence.
BU955838
ACCESSION
BU955838.1 GI:24185410
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 948)
REFERENCE
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLCM3052 row: e column: 01
High quality sequence stop: 592.
Location/Qualifiers

FEATURES

source

1..948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6729290"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH_MGC_107"
/note="Organ: breast; Vector: pOT87; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAGAG(G). Library constructed by
ling Hong in the laboratory of Gerald W. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 6.7%; Score 412.8; DB 5; Length 948;
Best Local Similarity 92.1%; Pred. No. 5.6e-26;
Matches 446; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

Oy		1	CACCATGGGCTCAGCGCAATTTATGAAAGCCAAATTACAGCCTTAAATATGAATGTGGAC	60
Db		43	CACCACGGGCTCGGCAATTTATGAAACCGAATATCACGCTTAAATATGAATGTGGAC	102
Oy		61	CTAAATATCCAGAAGAACCTCCCTTGTATAGATTTTGTAACAATAATATATGACTAGAG	120
Db		103	CTAATATATCCAGAAGACACCCCCCTTTGTATAGATTTTGTAACAATAATATATGACTAGAG	162
Oy		121	TTAATAGTTCCTAATGGAATGCTGAACCCAGAGCCATATCAGCGCTAGCAAATGGCAGA	180
Db		163	TAAATAGTTCCTAATGGAATGCTGGAGCCAGAGCCATATCAGTGCCTAGCAAATGGCAGA	222
Oy		181	ATTCAATATATCATCAAAAGTTAATCCCTTCAAGAGCTCACGCCCTATATGATGTCTAAAGAA	240
Db		223	ATTCAATATATGACATCAAAAGTTGTCTCTGCAAGAGCTTCGGCGCTTAATGTATCTAAAGAAA	282
Oy		241	ATGTGAACAGCGCTCAGGCATCTGTAAGAGACAGTGTTTACAGCAATTGATCAAAAAAGAAAA	300
Db		283	ATATGAACCTCCCTCAGCGCCGCCAGAGACAGTGTTCACGCAATTATCAAAAAAGAAAA	342
Oy		301	CCACAGGCGCTTCCCCTTCCCCTTCCCATATCTTGATGTGAAGCAGTCTTCATTTCCATATGAGT	360
Db		343	CCACAGGCGCTTCCCCTTCCCCTTCCCATATCGAATTTATATCAGTCTTCATTTCCACAGTAGT	402
Oy		361	AAATTTTCTAATATACAGTTGTATAGAGCTCCAAGTACTGGAAGAAAGAAAGCTCCCATTCAGAG	420
Db		403	AAATTTTCTAATATACAGTTGTATAGAGCTCCAAGTACTGGAAGAAAGAAAGCTCCCATTCAGAG	462
Oy		421	GAAATTTATCTTAAGATACAGTGTGAATGAATGACTA-TTTTTGTATCATTTGGAATATATAG	479
Db		463	GAAATTTATCTTAAGATACAGTGTGAATGAATGACTA-TTTTTGTATCATTTGGAATATATATAG	522
Oy		480	TTGT 483	
Db		523	TTGT 526	
RESULT 7				
B1258333			676 bp mRNA linear EST 17-JUL-2001	
DEFINITION			602966259P1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5107598 5',	
LOCUS			mRNA sequence.	
ACCESSION			B1258333	
VERSION			B1258333.1 GI:14814573	
KEYWORDS			EST.	
SOURCE			Homo sapiens (human)	
ORGANISM			Homo sapiens	
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
TITLE			1 (bases 1 to 676)	
JOURNAL			NIH-MGC http://mgc.nci.nih.gov/.	
COMMENT			National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs@email.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Incyte Genomics, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LLM11260 row: f column: 15 High quality sequence start: 8 Location/Qualifiers .676 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5107598" /tissue_type="cervical carcinoma cell line" /lab_host="DH10B"	
FEATURES				
Source				

Query Match	6.6*	Score 412;	DB 4;	Length 676;
Beat Local Similarity	92.58;	Pos. No. 8.2e-26;		
Matches 433;	Conservative 0;	Mismatches 35;	Indels 0;	Gaps 0;
<p> /clone_11b="NIH MGC 12" /inote="Organ: cervix; Vector: PCMV-SPORT6; Site 1: Nci1; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.4 kb. Library prepared by Life Technologies." </p>				
ORIGIN				
Query	16	CAATTGTGAAGCCAAATATATACAGCCTTAATAATAGATGTGACCTTAATACCCAGAG	75	
Db	177	CAATTTATGAAACCGAATATACGCTTAAATATGAAATGTGACCTTAATACCCAGAG	236	
Query	76	AACTCCCTTTGTAGATTGTAACAAAAATTAATATAGTAGAGTTAATAGTTCTAATG	135	
Db	237	CACCCCTTTGTAGATTGTAACAAAAATTAATATAGTAGAGTTAATAGTTCTAATG	296	
Query	136	GAATGTGMAACCAAGACCCATATCAGCCTAGCAAAATGGCAGATTCATATATCATCA	195	
Db	297	GAGTGTGMAACCAAGACCCATATCAGTGTAGCAAAATGGCAGAAATTCATATATGATCA	356	
Query	196	AAGTATTCCTCAAGAGCTTCAGGCGCTATATGATGTCTAAAGAAATGTGAAAGCCCTC	255	
Db	357	AAGTGTTCCTGCAAGAGCTTCGCGCTATATGATGTCTAAAGAAATATGAAATCTCCCTC	416	
Query	256	AGCCATCTGGAAGACAGTGTTCACAGCAATTGATCAAAAAGAAAACACAGGCCCTTCCC	315	
Db	417	AGCGCGCCGGAAGACAGTGTTCACAGCAATTATCAAAAAGAAAACACAGGCCCTTCCC	476	
Query	316	CTTCCCCCAATCTTGATGTAGACAGTCTTCATTTCCATATGATTAATTTCTAGATAC	375	
Db	477	CTTCCCCCAATCTTGATTAATCATGCTTCATATTTCCACAGTACTAATTTCTAGATAC	536	
Query	376	AGCTTTAGAGCTCAAGTACTGGAAGAAAGCCCATTCAAAGAAATTTATCTTAAG	435	
Db	537	GCTTTAGAGCTCAAGTACTGGAAGAAAGCCCATTCAAAGAAATTTATCTTAAG	596	
Query	436	ATACTGTAATGATCTAATTTTGTACATTTGGATATATTAAGTTGT	483	
Db	597	ATACTGTAATGATCTAATTTTGTCTCATTTGAAAATATATTAAGTTGT	644	
RESULT 8				
BUS95722	851 bp	mRNA	linear	EST 20-SEP-2002
LOCUS	AGENCOURT_8908446	NIH_MGC_142	Homo sapiens cDNA clone IMAGE:6451603	
DEFINITION	5', mRNA sequence.			
ACCESSION	BUS95722			
VERSION	BUS95722.1	GI:23247481		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
TITLE	1 (bases 1 to 851)			
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/.			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Unpublished (1999)			
	Contact: Robert Strausberg, Ph.D.			
	Email: cgabs-remail.nih.gov			
	Tissue Procurement: NCI			
	CDNA Library Preparation: Michael Brownstein Laboratory			
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)			
	DNA Sequencing by: Agencourt Bioscience Corporation			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LNL at:			
	http://image.lnl.gov			
	Plate: LICM624 row: b column: 20			
	High quality sequence stop: 455.			
	Location/Qualifiers			
	1..851			

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6451603"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_id="NIH_MGC_142"
/note="Vector: pDNR-Lib; Site 1: SfiI (ggccatcatggc); Site 2: SfiI (ggccgcctggcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%; and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:
5'-AAGCAGTGTATCAAGCAGAGTGGCCATTACGGCCGG-3' and 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the >0.5 kb size fraction (other fractions present in NIH_MGC_141). Library created in the laboratory of W. Brownstein (NIMH, NIH). Note: this is a NIH_MGC library."

ORIGIN

Query Match 6.6%; Score 411.8; DB 5; Length 851;
Best Local Similarity 91.9%; Pred. No. 7.3e-26;

Matches 445; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY 1 CACCATGGGCTCAGGCAATTATGAAAGCCAAATATACAGCTTAAATGAAATGTGAC 60
DB 30 CACCACGGGCTCGGACAAATTATGAAAGCCAAATATACAGCTTAAATGAAATGTGAC 89
QY 61 CTAATATACCGAAGAACTCCCTTGTAGATTGTGAATAAATAATTAATATGATGAG 120
DB 90 CTAATATACCGAAGAACTCCCTTGTAGATTGTGAATAAATAATTAATATGATGAG 149
QY 121 TTAATAGTTCTAATGAAATGTGAACCAAGACCATATACGCGCTAGCAAAATGGCAG 180
DB 150 TAAATAGTTCTAATGAAATGTGAACCAAGACCATATAGCTATGCAAAATGGCAG 209
QY 181 ATTCATATATCATCAAAATGATATCTTCAGAGCTTCAGCGCTTAATGATGTCTAAAGAA 240
DB 210 ATTCATATATGATCAAAAGTGTCTCTGCAAGAGCTTCGCGCTTAATGATGTCTAAAGAA 269
QY 241 ATGTGAAGCGCTCAGCCATCTGAAGAGAGGTGTACAGCAATGATCAAAAGAAAGAA 300
DB 270 ATATGAACTCTCTCAGCGCGCCGAGAGAGAGGTGTACAGCAATGATCAAAAGAAAGAA 329
QY 301 CCACAGGCGCTTCCCTCCCATCTTATGATGTAAGAGTCTTCATTTTCCATAGTAGT 360
DB 330 CCACAGGCGCTTCCCTCCCATCTTATGATGTAAGAGTCTTCATTTTCCACAGTAGT 389
QY 361 AAATTTTCTAGATACAGCTTGTAGAGCTCAAGATCTGAAAGAAAGCTCCATCAAG 420
DB 390 AAATTTTCTAGATACAGCTTGTAGAGCTCAAGATCTGAAAGAAAGCTCCATCAAG 449
QY 421 GAAATTTATCTAAGATCTGTAATGATCTAA-TTTTGTAGACTTTGGAATATATTAAG 479
DB 450 GAAATTTATCTAAGATCTGTAATGATCTAA-TTTTGTAGACTTTGGAATATATTAAG 509
QY 480 TTGT 483
DB 510 TTGT 513

RESULT 9
BE393340
LOCUS BE393340 727 bp mRNA linear EST 21-JUL-2000
DEFINITION 601311002F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:6332568 5',
ACCESSION BE393340
VERSION BE393340.1 GI:9338801

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:
<http://image.llnl.gov>
Plate: L1CM324 row: c column: 01
High quality sequence stop: 610.
Location/Qualifiers
1..727

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6332568"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_44"
/note="Organ: uterus; Vector: pOT87; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using 2AP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 6.6%; Score 411; DB 2; Length 727;
Best Local Similarity 92.5%; Pred. No. 9.4e-26;

Matches 432; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 16 CAATTTATGAAGCAAAATATACAGCTTAAATGAAATGTGAGCACTTAATACCGAGAG 75
DB 165 CAATTTATGAAGCAAAATATACAGCTTAAATGAAATGTGAGCACTTAATACCGAGAG 224
QY 76 AACTCCCTTGTAGATTGTGAACAAATTAATATGATGATGATGATGATGATG 135
DB 225 CACCCCTTGTAGATTGTGAACAAATTAATATGATGATGATGATGATGATG 284
QY 136 GAAATGTGAACCAAGACCATATACAGCTTGAAGAAATGGCAATTCATATCATCA 195
DB 285 GAGTGTGAGCCCAAGACCATATACAGCTTGAAGAAATGGCAATTCATATCATCA 344
QY 196 AAGTTATCTTCAAGAGCTTCAGCGCTTAATGATGTCTAAAGAAATGTGAAGCGCTC 255
DB 345 AAGTTATCTTCAAGAGAGCTTCAGCGCTTAAATGATGTCTAAAGAAATGTGAAGCGCTC 404
QY 256 AGCCATCTGAAGAGAGAGCTTGAAGCAATGATCAAAAGAAAGAAAGCAAGCGCTTCC 315
DB 405 AGCCCGCCGAAGAGAGAGCTTGAAGCAATGATCAAAAGAAAGAAAGCAAGCGCTTCC 464
QY 316 CTCCCGCATCTGATGATGAGAGTCTTCATTTTCCATAGATTAATTTCTAGATAC 375
DB 465 CTCCCGCATCTGATGATGATGAGAGTCTTCATTTTCCATAGATTAATTTCTAGATAC 524
QY 376 AGCTTGTAGAGCTCAAGATCTGGAAGAAAGAGCTCCATCAAGAAATTTATCTTAAG 435
DB 525 GTCTTGTAGAGCTCAAGATCTGGAAGAAAGAGCTCCATCAAGAAATTTATCTTAAG 584
QY 436 ATACTGTAAATGATCTAATTTTGTGATCAATTTGGAATATATTAAGTTG 482
DB 585 ATACTGTAAATGATCTAATTTTGTGATCAATTTGGAATATATTAAGTTG 631

RESULT 10
BI489733 708 bp mRNA linear EST 28-AUG-2001
LOCUS 60302230F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173398 5',
DEFINITION mRNA sequence.
ACCESSION BI489733
VERSION BI489733.1 GI:15328961
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 708)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LLM11431 row: 1 column: 07
High quality sequence stop: 704.
Location/Qualifiers
1..708
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5173398"
/lab_host="DH10B"
/clone.lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 6.6%; Score 410.8; DB 4; Length 708;
Best Local Similarity 92.1%; Pred. No. 1e-25;
Matches 444; Conservative 0; Mismatches 37; Indels 1; Gaps 1;
3 CCATGGGCTCAGGCAATTTATGAAGCCAAATATACAGCTTAATAATAGATGAGACT 62
1 CCAAGGCTCGACAAATTTATGAAGCCAAATATACAGCTTAATAATAGATGAGACT 60
63 AAATACCCAGAAAGAACTCCCTTTGTAGATTGTAAACAAATTAATATGATGAGATT 122
61 AAATACCCAGAAAGCAACCCCTTTGTAGATTGTAAACAAATTAATATGATGAGATT 120
123 AATAGTTCTTAATGAATGTGTAACCAAGAGCCATATACGCGCTAGCAAAATGGCAGAT 182
121 AATAGTTCTTAATGAATGTGTAACCAAGAGCCATATCAGTCTAGCAAAATGGCAGAT 180
183 TCATATATCATGAAGTTATCCTTCAAGAGCTTCAGCGCTAATGATGTCTAAAGAAAT 242
181 TCATATATCATGAAGTTATCCTTCAAGAGCTTCAGCGCTAATGATGTCTAAAGAAAT 240
243 GTGAAAAGCCCTCAGCCATCTGAAGAGCAGTGTACAGCAATGTATCAAAAGAAAACC 302
241 ATGAAAGCTCCCTCAGCGCCGAGAGAGCAGTGTACAGCAATGTATCAAAAGAAAACC 300

ORIGIN
Query Match 6.5%; Score 403.2; DB 5; Length 894;
Best Local Similarity 92.5%; Pred. No. 3.7e-25;
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;
16 CAATTATGAAGCCAAATATACAGCTTAATAATGAATGTGACCTTAATACCAAG 75
169 CAATTATGAAGAAACCAATATACAGCTTAATAATGAATGTGACCTTAATACCAAG 228
76 AACTCCCTTTGTAAAGATTGTGAACAAATTAATATAGTAACTAATGTTAATG 135
229 CACCCCTTTGTAAAGATTGTGAACAAATTAATATATGAATGAATAGTTCTAATG 288
136 GAATGTGAACCAAGAGCCATATCAGGCTAGCAAAATGCGAGAAATCATATATCATCA 195
289 GATGTGACCAAGAGCCATATCAGTCTAGCAAAATGCGAGAAATCATATATCATCA 348

RESULT 11
BUS07729 894 bp mRNA linear EST 12-SEP-2002
LOCUS AGENCOURT_10095740 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6501710
DEFINITION 5', mRNA sequence.
ACCESSION BUS07729
VERSION BUS07729.1 GI:22813962
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 894)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LLM14056 row: 1 column: 15
High quality sequence stop: 621.
Location/Qualifiers
1..894
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6501710"
/issue_type="telomysarcoma"
/lab_host="DH10B (phage-resistant)"
/clone.lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

ORIGIN
Query Match 6.5%; Score 403.2; DB 5; Length 894;
Best Local Similarity 92.5%; Pred. No. 3.7e-25;
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;
16 CAATTATGAAGCCAAATATACAGCTTAATAATGAATGTGACCTTAATACCAAG 75
169 CAATTATGAAGAAACCAATATACAGCTTAATAATGAATGTGACCTTAATACCAAG 228
76 AACTCCCTTTGTAAAGATTGTGAACAAATTAATATAGTAACTAATGTTAATG 135
229 CACCCCTTTGTAAAGATTGTGAACAAATTAATATATGAATGAATAGTTCTAATG 288
136 GAATGTGAACCAAGAGCCATATCAGGCTAGCAAAATGCGAGAAATCATATATCATCA 195
289 GATGTGACCAAGAGCCATATCAGTCTAGCAAAATGCGAGAAATCATATATCATCA 348

QY 196 AAGTATCCCTTCAAGAGCTTCAAGCGGCTTAATGATGTCCTAAGAAATGTGAAACGGCCCTC 255
 DB 349 AAGTTGCTCCGAGAGAGCTTCGGCGCTTAATGATGTCCTAAGAAATGTGAAACGGCCCTC 408
 QY 256 AGCCATCTGAAGACAGTGTACAGCAATTCATCAAAAAAACCACAGGCCCTTCCC 315
 DB 409 AGCCGCCCAAGACAGTGTACAGCAATTCATCAAAAAAACCACAGGCCCTTCCC 468
 QY 316 CTTCCGCCCAATCTGAAGAGCTTCAATTTCCATAGTAGTAATTTTCTAGATAC 375
 DB 469 CTTCCGCCCAATCTGAAGAGCTTCAATTTCCATAGTAGTAATTTTCTAGATAC 528
 QY 376 AGCTTGTAGAGCTCAAGATCTGGAAGAAAGAGTCCCATTCAGAAAGAAATTTATCTTAAG 435
 DB 529 GTCTGTAGACCTCAAGATCTGGAAGAAAGAGTCCCATTCAGAAAGAAATTTATCTTAAG 588
 QY 436 ATACTGTAAATGATTAATA-TTTTGTACATTTGGAATTAATAATGTTGT 483
 DB 589 ATACTGTAAATGATTAATAATTTTGTGTCAATTTGGAATTAATAATGTTGT 637

RESULT 12
 BC005215 1180 bp mRNA 1linear HTC 19-NOV-2003
 LOCUS Homo sapiens ubiquitin-conjugating enzyme E2 variant 1, mRNA (cDNA
 clone IMAGE:3996506), containing frame-shift errors.
 ACCESSION BC005215
 VERSION BC005215.1 GI:14709942
 KEYWORDS HTC.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1180)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Dege, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, D., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stappleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullan, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richardson, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huylk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Sanchez, J., Helton, E., Ketterman, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E.,
 Schneringer, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL MEDLINE
 PUBLISHED 22388257
 REFERENCES 12477932
 2 (bases 1 to 1180)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (27-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Center, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: gcabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMUN)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mc@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILMUN at: <http://image.llnl.gov>
 Series: IRL Plate: 16 Row: 9 Column: 24
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 15718758
 This clone has the following problem: frame shifted.

FEATURES

source
 1..1180
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3996506"
 /tissue_type="Bladder, carcinoma"
 /clone_id="NIH_MGC_53"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"

ORIGIN

Query Match 6.5%; Score 403; DB 3; Length 1180;
 Best Local Similarity 92.4%; Pred. No. 3,2e-25;
 Matches 435; Conservative 0; Mismatches 35; Indels 1; Gaps 1;
 QY 14 GGCATTTATGAAAGCCAAATATACAGCTTAAATGAGAGCTTAATACCCAGA 73
 DB 674 GACATTTATGAAAGCCAAATATACAGCTTAAATGAGAGCTTAATACCCAGA 733
 QY 74 AGAATCTCCCTTGTGAATTTGTAAACAAATTAATGATAGTAAATAGTTCTTA 133
 DB 734 AGCACCCTCTTGTGAATTTGTGAACAAATTAATGAGAGCTTAATAGTTCTTA 793
 QY 134 TCGAATGTGAACCCAGAGCCATATCAGCGCTTACAAATGCGAATTCATATCAT 193
 DB 794 TCGAATGTGAACCCAGAGCCATATCAGCGCTTACAAATGCGAATTCATATCAT 853
 QY 194 CAAAGTTATCTCTTCAAGAGCTTCAAGCGCTTATGATGCTTAAGAAATGTGAAACGCC 253
 DB 854 CAAAGTTGCTCGCAAGAGCTTCGCGCTTATGATGCTTAAGAAATGTGAAACGCC 913
 QY 254 TCAGCATCTGAAGACAGTGTACAGCAATGATCAAAAGAAACCAAGAGCCCTTC 313
 DB 914 TCAGCATCTGAAGACAGTGTACAGCAATGATCAAAAGAAACCAAGAGCCCTTC 973
 QY 314 CCGTTCGCCCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 373
 DB 974 CCGTTCGCCCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1033
 QY 374 ACAAGTTGAGAGCTCAAGATCTGGAAGAAAGTCCCATTCAGAAATTTATCTTA 433
 DB 1034 ACAGTCTGTGAGAGCTCAAGATCTGGAAGAAAGTCCCATTCAGAAATTTATCTTA 1093
 QY 434 AGATCTGTAATGATTAATA-TTTTGTACATTTGGAATTAATAATGTTGT 483
 DB 1094 AGATCTGTAATGATTAATAATTTTGTGTCAATTTGGAATTAATAATGTTGT 1144

RESULT 13
 CN346858 489 bp mRNA 1linear EST 16-MAY-2004
 LOCUS CN346858
 DEFINITION 17000532816012 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
 ACCESSION CN346858
 VERSION CN346858.1 GI:47346792
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 489)
 AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Muraige, J., Flek, G.J.,

Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M. S., Mandalam, R., Lebkowksi, J. and Stanton, L. W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Gen Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@gen.com
Insert length: 489 Std Error: 0.00.
Location/Qualifiers

FEATURES
Source

1. 489
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_1ib="GRN_EB"
/note="oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowth derived from h9s cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN

Query Match 6.5%; Score 402.6; DB 7; Length 489;
Best Local Similarity 92.5%; Pred. No. 6.1e-25;
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

16 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGTGGACCTAAATACCGAAG 75
2 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGTGGACCTAAATACCGAAG 61
76 AACTCCCTTTGTAAGATTGTGAACAAAATTAATAGTGAAGTAAATAGTTCTAATG 135
62 CACCCCTTTGTAAGATTGTGAACAAAATTAATAGTGAAGTAAATAGTTCTAATG 121
136 GAATGTGAAGCCAAAGACCATATCAGCGCTAGCAAAATGCGAATTCATATATATCA 195
122 GAGTGTGAAGCCAAAGACCATATCAGCGCTAGCAAAATGCGAATTCATATATATCA 181
136 AAGTTATCTTCAAGAGCTTCAGCGCTAATGATGTCTTAAGAAATGTGAAGCCCTC 255
182 AAGTTATCTTCAAGAGCTTCAGCGCTAATGATGTCTTAAGAAATGTGAAGCCCTC 241
256 AAGCATCTGAAGAGAGCTTACAGCAATTCATCAAAAAGAAACACAGGCGCTTCC 315
242 AAGCATCTGAAGAGAGCTTACAGCAATTCATCAAAAAGAAACACAGGCGCTTCC 301
316 CTTCCCTTCTTCAAGAGCTTCATTCATTCATAGTAAATTTTCTAGATAC 375
302 CTTCCCTTCTTCAAGAGCTTCATTCATTCATAGTAAATTTTCTAGATAC 361
376 AAGTTATCTTCAAGAGCTTCAGCGCTAATGATGTCTTAAGAAATGTGAAGCCCTC 435
362 GCTTGTGAAGCTCAAGAGAGCGAAGAGAGCTCCATTCAGAAAGAAATTTATCTTAAG 421
436 AATCTGTAAATGATATCTAA-TTTTGTAGATTTGGAATTAATTAAGTTGT 483
422 AATCTGTAAATGATATCTAA-TTTTGTGCAATTTGGAATTAATTAAGTTGT 470

RESULT 14
AI376271/c 491 bp mRNA linear EST 21-JAN-1999
LOCUS AI376271
DEFINITION CASH12.x1 Soares fetal_Nb2HP8.9w Homo sapiens cDNA clone
IMAGE:2048039 3', similar to TR:Q13403 Q13403 DNA-BINDING PROTEIN.
// mRNA sequence.
ACCESSION AI376271
VERSION AI376271.1 GI:4176261
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 491)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.jnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 398.
Location/Qualifiers

FEATURES
Source

1. 491
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2048039"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/clone_1ib="Soares fetal_fetus_Nb2HP8.9w"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCTTATTTTCTTTTCTTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73D vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 6.5%; Score 402.6; DB 1; Length 491;
Best Local Similarity 92.5%; Pred. No. 6.1e-25;
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

16 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGTGGACCTAAATACCGAAG 75
481 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGTGGACCTAAATACCGAAG 422
76 AACTCCCTTTGTAAGATTGTGAACAAAATTAATAGTGAAGTAAATAGTTCTAATG 135
421 CACCCCTTTGTAAGATTGTGAACAAAATTAATAGTGAAGTAAATAGTTCTAATG 362
136 GAATGTGAAGCCAAAGACCATATCAGCGCTAGCAAAATGCGAATTCATATATATCA 195
361 GAGTGTGAAGCCAAAGACCATATCAGCGCTAGCAAAATGCGAATTCATATATATCA 302
196 AAGTTATCTTCAAGAGCTTCAGCGCTAATGATGTCTTAAGAAATGTGAAGCCCTC 255
301 AAGTTATCTTCAAGAGCTTCAGCGCTAATGATGTCTTAAGAAATGTGAAGCCCTC 242
376 AAGTTATCTTCAAGAGCTTCAGCGCTAATGATGTCTTAAGAAATGTGAAGCCCTC 435
241 AAGCATCTGAAGAGAGCTTACAGCAATTCATCAAAAAGAAACACAGGCGCTTCC 182
316 CTTCCCTTCTTCAAGAGCTTCATTCATTCATAGTAAATTTTCTAGATAC 375
181 CTTCCCTTCTTCAAGAGCTTCATTCATTCATAGTAAATTTTCTAGATAC 122
436 AATCTGTAAATGATATCTAA-TTTTGTAGATTTGGAATTAATTAAGTTGT 483
422 AATCTGTAAATGATATCTAA-TTTTGTGCAATTTGGAATTAATTAAGTTGT 470

RESULT 15
BM739173 500 bp mRNA linear EST 01-MAR-2002
LOCUS K-EST0008565 S3SNUI6 Homo sapiens cDNA clone S3SNUI6-3-H01 5', mRNA
DEFINITION
ACCESSION BM739173
VERSION BM739173
KEYWORDS EST, GI:19060502
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 500)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.U., Cheong,J.E., Sohn,H.Y., Kim,U.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongseung@mail.kribb.re.kr
Plate: 3 row: H column: 01
High quality sequence stop: 500.
Location/Qualifiers
1. 500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S3SNUI6-3-H01"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/lab_host="T1010F"
/clone_1ib="S3SNUI6"
/note="Organ: Stomach; Vector: pTZ19RP1; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli 'Top10' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN
Query Match 6.5%; Score 402.6; DB 4; Length 500;
Best Local Similarity 92.5%; Pred. No. 6e-25;
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;
QY 16 CAATTATGAAAGCCAAATATACAGCCTTAAATAGATGTGACCTAAATACCCAGAAG 75
DB 16 CAATTATGAAAGCCAAATATACAGCCTTAAATAGATGTGACCTAAATACCCAGAAG 75
QY 76 AACTCCCTTTGTAGATTGTGTAACAAAATTAATATGATGAGTTAAATAGTTCTTAATG 135
DB 76 CAACCCCTTTGTAGATTGTGTAACAAAATTAATATGATGAGTTAAATAGTTCTTAATG 135
QY 136 GAATGTGAACCCAAAGCCATATCAGCGCTAGCAAAATGGCAGAAATCATATATCATCA 195
DB 136 GAGTGTGAGCCCAAGCCATATCAGCGCTAGCAAAATGGCAGAAATCATATATCATCA 195

QY 196 AAGTATCTCTTCAAGAGCTTCAAGCGCTTAATGATGTCTTAAGAAATGTGAAACGCCCTTC 255
DB 196 AAGTGTCTCTTCAAGAGAGCTTTCGCGCCTTAATGATGTCTTAAGAAATATAGAAATCTCCCTC 255
QY 256 AGCCATCTGAAGAGACAGTGTTCACAGCAATTGATCAAAAAAGAAAAACACAGGCGCTTCCC 315
DB 256 AGCCGCGCGAAGAGACAGTGTTCACAGCAATTGATCAAAAAAGAAAAACACAGGCGCTTCCC 315
QY 316 CTTCGCCCATCTTGTAGTGAAGAGAGTCTTCATTTTCCATAGTAGTAAATTTTCTAGATAC 375
DB 316 CTTCGCCCAATTCGATTTAATCAAGTCTTCATTTTCCACAGTAGTAAATTTTCTAGATAC 375
QY 376 AGCTGTGAAGTCCAAAGTACAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 435
DB 376 GTCTGTGAAGTCCAAAGTACAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 435
QY 436 ATACTGTAAATGATGACTTAA-TTTTGTACATTTGGAAATATATAGTTGT 483
DB 436 ATACTGTAAATGATGACTTAAATTTTGTCCATTTGGAATATATAGTTGT 484

Search completed: December 27, 2004, 10:56:48
Job time : 11864.7 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using ew model

Run on: December 26, 2004, 17:40:11 ; Search time 9563.51 Seconds
(without alignment)

19051.434 Million cell updates/sec

Title: US-10-786-065-3_COPY_1_5000

Perfect score: 5000

Sequence: 1 tataggccatgcgtgcgtc.....ctatcttcttctatcttgc 5000

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 32822875 seqs, 1821986508 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Database : Listing first 45 summaries

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_est4: *
5: gb_est5: *
6: gb_est6: *
7: gb_est7: *
8: gb_est8: *
9: gb_est9: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	473.8	9.5	856	9	CLJ01570
C 2	440.2	8.8	483	8	AO138494
C 3	433	8.7	819	8	BZ879668
C 4	402.8	8.1	482	8	AO466201
C 5	401.4	8.0	431	8	AO052162
C 6	374.8	7.5	436	8	B34439
C 7	359.4	7.2	676	8	BZ879665
C 8	325.2	6.5	5940	3	CR627133
C 9	316.2	6.3	472	8	AO709562
C 10	302.8	6.1	6146	8	AO839831
C 11	293.2	5.9	1300	8	AF057104
C 12	292.8	5.9	321	8	AO072178
C 13	291.6	5.8	1279	3	BC037921
C 14	288.2	5.7	647	4	BM559057
C 15	284.8	5.7	660	4	AL120269
C 16	282	5.6	564	5	BU860291
C 17	282	5.6	566	5	BU957747
C 18	281.2	5.6	1206	3	CR619443
C 19	279.2	5.6	515	8	AO019249
C 20	279	5.6	1644	3	BC037923
C 21	277	5.5	2100	3	AF289610
C 22	276.4	5.5	669	3	AG155696
C 23	276.4	5.5	714	6	CD366358
C 24	276	5.5	576	6	CA427630

25	275.8	5.5	444	8	B89781
26	275.8	5.5	758	5	BQ708090
27	275.2	5.5	1029	4	BM471041
28	274.6	5.5	513	5	BM642735
C 29	274.4	5.5	3288	3	HS802613
C 30	274.2	5.5	3713	3	BC041839
C 31	274	5.5	924	5	BM674842
C 32	273.8	5.5	709	6	CD359469
C 33	273.8	5.5	736	7	CN280511
C 34	273.8	5.5	772	5	BU928605
C 35	273.8	5.5	882	5	BQ227019
C 36	273.4	5.5	701	5	BU615970
C 37	273.2	5.5	3863	5	BC035034
C 38	273	5.5	664	8	AO343449
C 39	272.8	5.5	657	9	AG182991
C 40	272.6	5.5	822	4	BG249643
C 41	272.2	5.4	514	8	AO002001
C 42	272.2	5.4	541	6	CA437967
C 43	272.2	5.4	727	6	CA440514
C 44	272.2	5.4	1050	6	CD252308
C 45	272	5.4	471	8	AZ695036

ALIGNMENTS

RESULT 1	CLJ01570/c	856 bp	DNA	linear	GSS 20-FEB-2004
LOCUS	gb608 CH257	PP0080Q4	SP6	A08CD726	CHORI-257
DEFINITION	Microcebus murinus genomic clone CH257-8B16,	genomic survey			
ACCESSION	CLJ01570				
VERSION	CLJ01570.1	GI:42718759			
KEYWORDS	GSS.				
SOURCE	Microcebus murinus (gray mouse lemur)				
ORGANISM	Microcebus murinus				
REFERENCE	1 (bases 1 to 856)				
AUTHORS	Dore, C., Zhang, X., McKee, K.H., Forgetta, V., Lepage, P., Hudson, T.J.				
TITLE	MUGQ Lemur Physical Mapping Project				
JOURNAL	Unpublished (2004)				
COMMENT	Other GSSes: gb608_CH257_PP0080Q4_T7u_A08CD726				
	Contact: Dewar K				
	McGill University and Genome Quebec Innovation Centre				
	740 Dr. Penfield Room 7214, Montreal, QC, Canada, H3H 1A4				
	Tel: 514 398 3311 x00089				
	Fax: 514 398 1795				
	Email: ken.dewar@mcgill.ca				
	Plate: 8 row: B column: 16				
	Seq primer: SP6 : CTGGCCGTCGACATTAGG				
	Class: BAC ends.				
FEATURES	Location/Qualifiers				
source	1..856				
	/organism="Microcebus murinus"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:30608"				
	/clone="CH257-8B16"				
	/sex="female"				
	/cell_type="kidney"				
	/clone_lib="CHORI-257 Lemur BAC Library"				
	/note="Vector: pTARBA2.1 vector; Site 1: EcoRI; Site 2: EcoRI. The CHORI-257 mouse lemur BAC library has been constructed by Michael Medvedov in Pieter de Jong's laboratory at BACPAC Resources, Children's Hospital Oakland Research Institute. The preparation in our laboratory followed the cloning approach developed in our laboratory (Osoegawa et al., 1998). Kidney was obtained from an adult female mouse lemur (Microcebus murinus) through John				

Allman, Division of Biology 216-76, Caltech, Pasadena, CA 91125. The local identification number for this animal is #8. A skin biopsy and a blood sample have been sent to the Coriell Cell Repositories to establish cell lines as part of the Integrated Primate Biomaterials and Information Resource (IPBIR) initiative. High-molecular-weight DNA was isolated from a kidney cell suspension prepared in a Dounce Homogenizer and the cells were embedded in agarose to stabilize the DNA. The embedded DNA was partially digested with a combination of EcoRI restriction enzyme and EcoRI methylase and size fractionated by pulsed-field electrophoresis. DNA fragments from the appropriate size fraction were cloned into the pTRBAC2.1 vector between the two EcoRI sites. The ligation products were transformed into DH10B (T1 resistant) electro-competent cells (Invitrogen). The library has been arrayed into 576 '384-well' microtiter dishes and also gridded onto twelve 22x22cm nylon high-density filters for screening by probe hybridization. Each hybridization membrane represents about 18,000 distinct mouse Lemur BAC clones, stamped in duplicate. Library characterization was performed by Qing Cao, Teresa Ren and Kazuhiro Osoegawa. This work was funded by NIH grant HG025323-01."

ORIGIN

Query Match 9.5%; Score 473.8; DB 9; Length 856;

Best Local Similarity 80.9%; Pred. No. 8.5e-48;

Matches 614; Conservative 0; Mismatches 137; Indels 8; Gaps 5;

```

QY 3378 ATCGCAGCTTACAGAGGAGTATGAGCAGTATGACATAGCTGGAGAGCTTTAGA 3437
DB 856 ATCTCTACAGTTACAAAGGATTTGACAGTGTAGTACTTTAGGTAGAGAGCTTTAGA 797
QY 3438 GATGAGATGCCGC-CCCCCAATTCATATTAAGCCAGGTGAGATCATAGAACT 3496
DB 796 GATGAGAGGTCCACGCCCTTTATTTTACATTAATAAGCCAGGAGACATCAGAGAGT 727
QY 3497 TCATACGACTCAGAGCCTGTGCAAGACATATGCGGACAGGAGAGACATGATTAAT 3556
DB 736 TCACAGATTCAGAGCCTGTGCAAGACATCCTGCAATAGGAGAGAGAGGATGTT 677
QY 3557 TAAACAGCCTTGAAGAAACCAACCTGCTGCTTAATTAATTCAGCCACTTAA 3616
DB 676 TAAATAGCCTGGAAGAAACCAACTG-CTTGCCCTAATTAATTCAGCAATCTGAAA 618
QY 3617 TGTTCATCAGCCTTTCCCTTCCTTCATTCAGAAATTCAGAAATAGACATTC 3676
DB 617 TATTTATAGCCTTTCCCTTCCTTCATTCAGAAATTCAGAAATAGACATTC 558
QY 3677 TCTACTACGACCCAAAGAAACATTCATCTTCAGGCGCTGTGGAGGACAGCTTGTA 3736
DB 557 TTTACTACTGACCCAAAGAAATTCATCTTCAGGCGGTGGAGGAGGACAGCTTGTA 498
QY 3737 AAGCGTCTTAACAGGTTTTTATATCCCTCCCTTAATCAATGACAGAGTTTGTAA 3796
DB 497 AGGTATCTCTAAAGATTTTATATATCCCTCCCTAATCATATGACAGAGTTTCAAT 438
QY 3797 GGAACCTGGAATTTGCTGCTTATTCCTTCACCTGAGCTTTATAGAAGAACTGAAGT 3856
DB 437 GGAACCTGGAATTTGCTGCTTATTCCTTCACCTGAGCTTTATAGAAGAACTGAAGT 378
QY 3857 GGTTCGCAAAATTAATGATCAAAAGATGATAATCCATGATTTTTTAAATTTGA 3916
DB 377 GGTTCGCAAAATTAATGATCAAAAGATGATAATCCATGATTTTTTAAATTTGA 322
QY 3917 AAATACACAAATGTGTGAGATTAATAATCTGTTATCAAAAGCTAGTACTAATTT 3976
DB 321 AAATACACAAATGTGTGAGATTAATAATCTGTTATCAAAAGCTAGTACTAATTT 262
QY 3977 TGGTAAACAACTGTTGTTAATATATGTAAGAAATCAATGATTT-CCCTTTTAACT 4034
DB 261 TGGTAAACAACTGTTGTTAATATATATGTAAGAAATCAATGATTTCACTTTTAACT 202

```

```

QY 4035 CAAGTGGGAAAGTGTGATGTGCTTTTCTTTATGTTACTCCAAATAGAGAAAGT 4094
DB 201 TAAGGTGGGAAATTAAGAGCTTTATTTATTTATTTATTTATTTATTTATTTAT 142
QY 4095 AATGCTCAATGTGTTGTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 4133
DB 141 AATGCTCAATGTGTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 103

```

RESULT 2

LOCUS A0138494 483 bp DNA linear GSS 24-SEP-1998

DEFINITION HS_3076_B1_B01_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3076 Col=1 Row=D, genomic survey

sequence.

ACCESSION A0138494

VERSION A0138494.1 GI:35292147

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 483)

REFERENCE Mahairas G.G., Wallace J.C., Smith K., Swartzell S., Holzman T., Hood L., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and

Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and

Hood L. Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

JOURNAL MEDLINE

PUBMED

COMMENT

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 3076 row: D column: 1

Class: BAC ends

High quality sequence strop: 483.

location/Qualifiers

1..483

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clone="Plate=3076 Col=1 Row=D"

/sex="male"

/clone_id="CIT Approved Human Genomic Sperm Library D"

/note="Organ: sperm; Vector: pBelOAc11; BAC Clones in E-Coli DH10B"

ORIGIN

Query Match 8.8%; Score 440.2; DB 8; Length 483;

Best Local Similarity 99.3%; Pred. No. 1.2e-43;

Matches 442; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 3665 AAATAGACATCTCTACTACTGACCCAAAGAAACATTTATCACTTCAGGCGCTGTGGAG 3724
DB 39 AAATAGACATCTCTACTACTGACCCAAAGAAACATTTATCACTTCCTTTAAGCCTGTGGAG 98
QY 3725 GCAAGTTGTGAAGCGTCTCTACAGGTTTTTATATTCCTCCCTAATATCAATGACA 3784
DB 99 GCAAGTTGTGAAGCGTCTCTACAGGTTTTTATATTCCTCCCTAATATCAATGACA 158
QY 3785 GAGTTTGTGAATGGAACCTGGAATTTGCTGCTTCATTCCTCCAGCCTGCTTATAGA 3844
DB 159 GAGTTTGTGAATGGAACCTGGAATTTGCTGCTTCATTCCTCCAGCCTGCTTATAGA 218
QY 3845 GAAACTGAAGTTGTTCTGCAAAATTAATGATGACAAAGATGATAATCTTGAATTT 3904
DB 219 GAAACTGAAGTTGTTCTGCAAAATTAATGATGACAAAGATGATAATCTTGAATTT 278

```

QY 3905 TTATATTTGGCAAAATACCAAAATGTCGAGAAATPAAATATCTCTTATCCAAAGCT 3964
DB 279 TTTATATTTGGCAAAATACCAAAATGTCGAGAAATPAAATATCTCTTATCCAAAGCT 338
QY 3965 AAGTACTAATTTTGGTAAACAACCACTTGTATTAATATATGTAAGATCCATGAAATC 4024
DB 339 AAGTACTAATTTTGGTAAACAACCACTTGTATTAATATATGTAAGATCCATGAAATC 398
QY 4025 CCCTTTAGTCAGAGTGGGAAAGTGGATGTCGCTTTTCTTTATGTTACTCAATAG 4084
DB 399 CCCTTTAGTCAGAGTGGGAAAGTGGATGTCGCTTTTCTTTATGTTACTCAATAG 458
QY 4085 AGAGAAAGTATGCTCAATAGTG 4109
DB 459 AGAGAAAGTATGCTCAATAGTG 483

RESULT 3
B2879668/c 819 bp DNA linear GSS 18-MAR-2003
LOCUS CH240_195A22.TVB CHORI-240 Bos taurus genomic clone CH240_195A22,
DEFINITION genomic survey sequence.
ACCESSION B2879668
VERSION B2879668.1 GI:29107070
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
1 (bases 1 to 819)
Zhao,S., Shetty,J., Shateman,S., Tsegaye,G., Geer,K.,
Svartbejrn,A., Gebregeorgis,E., Chen,D., Riggs,F., de Jong,P.,
Crawford,A.M. and McEwan,J.C.
Bovine BAC End Sequences from Library CHORI-240
Unpublished (2003)
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdjong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBMC) by AgResearch Ltd., New Zealand and The
Institute of Genomic Research (TIGR), USA.
Plate: 195 row: A column: 22
Seq primer: Sp6
Class: BAC ends.

FEATURES
source
1..819
Location/Qualifiers
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_195A22"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull H Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

ORIGIN
Query Match 8.7%; Score 433; DB 8; Length 819;
Best Local Similarity 78.9%; Pred. No. 76-43;
Matches 651; Conservative 0; Mismatches 140; Indels 34; Gaps 10;

QY 2202 CCGAGAGTACGCTGAGACCAAGAGCTGCTTCAAGATTTGTATCCAGAGAGA 2261
DB 819 CAGAGAGAGTCCCTGTGTATGACGAGAGCTGGATTCAGATTCATATCCAG----GA 764
QY 2262 GCATCTCTCTATTTGATTAACCAAGAGTTCAGACATCCCTTTTGTAGGGGATCTG 2321
DB 763 GCATCTCCCTCTATCTGATTAACCA--GAGTTCAGAGCTCCCTTTCTGTAGAGGGTCTG 707
QY 2322 ATCTCTCCGAGTGTCTTAAACCAATTAATGAAATTTCTATTAAAGTCACAGAAAT 2381
DB 706 CTTTCTCTGATTTAGACATTAAGCAATTAAGAAATTTCTATTAAAGTCACAGAAATG 647
QY 2382 TATGCTGTAGTATCAAAATTTGGGGAATTTCTGTAAACCAAGGAAATTAATCT 2441
DB 646 TGTGCTGTGTATCAAAATTCGGGGAATTTCTTTTAAACCAAGGGAATTAATCCCT 587
QY 2442 TGGCTTTGGGCTGCAGCAAACTCACTTGGCTTGAAGTGCAGAAAGTATCTCAAAAT 2501
DB 586 ATGCTTTGGGCTGCAGCAAACTCACTTGGCTTGAAGTGCAGAAAGTATCTTGAAT 527
QY 2502 CTCTAAGTCTTAATTTAGAGAGTGAATTTAAAGCAAGCTGAGTATTAGTTGTA 2561
DB 526 CTCTGAAGTACGAATTTAGCGAGTGAACCTCAAAACAGAGCTGGGAT---TTGTA 470
QY 2562 TGTATGATTTGAACCTTTAG--TAATTAATTCATGATTTATGCAATCCATAGA---- 2616
DB 469 TCTATGATTTGAACCTTTAG--TAATTAATTCATGATTTATGCAATCCATAGA 410
QY 2617 --TTATTCCTTCACGCAATTAATTTAGAGCATGAATAAAGGCAAGCTTAAGA 2674
DB 409 AGTTATTCCTTCACGCAATTAATTTAGAGCATGAATAAAGGCAAGCTTAAGA 350
QY 2675 GAAAAAGTTGCAAAACATCAACAAAGATTTAGTTAGCTGAATG--TAAGACAC 2733
DB 349 GAAAA--TTGGAACCATCAACAAAGATTTAGTTAGCTGAATG--TAAGACAC 292
QY 2734 A--TTTATGCTGAAGAAAGATTTAGTTAGCTGAATG--TAAGACAC 2791
DB 291 ATTTTATGCTGAAGAAAGATTTAGTTAGCTGAATG--TAAGACAC 232
QY 2792 -----CAGGATCTTCAAGAAAGTCCCTTTGGGGGTACAGAAAGCTTAGA 2840
DB 231 TTTTAAAAAGAACTGATGTTCAAGAAAGTACATTTGAAAGTGGAAAGCTTAGA 172
QY 2841 AACATTTGAAGTGAATTAAGAGCAATTAAGAAATTAAGTTTATCCAGGACATGAAT 2900
DB 171 AACATTTGAAGTGAATTAAGAGCAATTAAGAAATTAAGTTTATCCAGGACATGAAT 113
QY 2901 CTTTACTTGCATTAATTTTATTTTCTGCTCTTTTCTTTTCTAGCTAACAGACTTAA 2960
DB 112 CTTGACTTGCATTAATTTTATTTTCTGCTCTTTTCTTTTCTAGCTAACAGACTTAA 53
QY 2961 AGAAGCATGATTTTCATGATTTTCACTTCACTTCACTTCACTTCACTTCACTTCA 3005
DB 52 AGAAGCATGATTTTCATGATTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 8

RESULT 4
AQ466201/c 482 bp DNA linear GSS 23-APR-1999
LOCUS HS_5157_A2_B06_SPEE RPCI-11 Human Male BAC library Homo sapiens
DEFINITION genomic clone Plate=73 Col=12 Row=C, genomic survey sequence.
ACCESSION AQ466201
VERSION AQ466201.1 GI:4643296
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 482)
Mahlira,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and

TITLE
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE
99380589
PubMed
10449764
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
Plate: 733 row: C column: 12
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 482.

FEATURES

source
1..482
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=733 Col=12 Row=C"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"

ORIGIN

Query Match 8.1%; Score 402.8; DB 8; Length 482;
Best Local Similarity 91.5%; Pred. No. 3.7e-39;
Matches 439; Conservative 0; Mismatches 37; Indels 4; Gaps 1;
QY 4301 ATGCTGTATTTATTTTACTTGTAAACAGATTGGAGGACATACATTTATT 4360
DB 481 ATGCTGTATTTATTTTACTTGTAAACAGATTGGAGGACATACATTTATT 422
QY 4361 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4416
DB 421 ATATCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 362
QY 4417 ACCTGTAGTGCAGTGTGAGATCTGAGATGACGACCTCTGCTGAGTTCAAGC 4476
DB 361 AGCTGTAGTGCAGTGTGAGATCTGAGATGACGACCTCTGCTGAGTTCAAGC 302
QY 4477 AACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4536
DB 301 AACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242
QY 4537 ACTTTTGAATGCGAGAAACATGCGAATTAAGGCCACATTTATTTATTTATTTATTT 4596
DB 241 ACTTTTGAATGCGAGAAACATGCGAATTAAGGCCACATTTATTTATTTATTTATTT 182
QY 4597 TCTCATATCAATTTGGCATGGTAACTGTTTCAAGTGGCTGGAGCGGGGGACCCCT 4656
DB 181 TCTCATATCAATTTGGCATGGTAACTGTTTCAAGTGGCTGGAGCGGGGGACCCAT 122
QY 4657 GGAACATCTTGGATTCATGGGACCATGAGACATTTCTGATCCTCTCTTCTGAGTTCTGA 4716
DB 121 GGAACATCTTGGATTCATGGGACCATGAGACATTTCTGATCCTCTCTTCTGAGTTCTGA 62
QY 4717 CTTTATTTGTTCTGACAGACCTTTTCCAGCCGGAAGTTTACAGAAATTCATTTATTTCTTT 4776
DB 61 CTTTATTTGTTCTGACAGACCTTTTCCAGCCGGAAGTTTACAGAAATTCATTTATTTCTTT 2

RESULT 5
AA052162
LOCUS
DEFINITION
RPC111-50K8.TK RPCI-11 Homo sapiens genomic clone RPCI-11-50K8,
genomic survey sequence.
ACCESSION
AA052162
VERSION
AA052162.1 GI:3349199
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 431)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Baas,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
Venter,J.C.
Use of BAC End Sequences for Sequence-Ready Map Building (1998)
Unpublished (1998)
Other GSSs: RPC111-50K8.TJ
COMMENT
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamad@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Class: BAC ends.

FEATURES

source
1..431
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/db_xref="GDB:7519063"
/clone="Plate=733 Col=12 Row=C"
/clone_lib="RPCI-11-50K8"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI; RPC111 Human Male BAC Library"

ORIGIN

Query Match 8.0%; Score 401.4; DB 8; Length 431;
Best Local Similarity 98.5%; Pred. No. 5.7e-39;
Matches 405; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3647 ATTGAGAAATTCAGAAATTAAGATCTCTACTGACCCGAAAGAAATTAATAC 3706
DB 21 ATCGCGGAATTCAGAAATTAAGATCTCTACTGACCCGAAAGAAATTAATAC 80
QY 3707 TCTTAGGCTGTGGAGGACAGTTGTTAAAGGCTCTTACAGAGTTTATTAATCCCT 3766
DB 81 TCTTCAGGCTGTGGAGGACAGTTGTTAAAGGCTCTTACAGAGTTTATTAATCCCT 140
QY 3767 CCTTAATCAAGACAGAGTTTGTAAATGCAACCGGAATTTGCTGCTTCAATTCCTC 3826
DB 141 CCTTAATCAAGACAGAGTTTGTAAATGCAACCGGAATTTGCTGCTTCAATTCCTC 200
QY 3827 CACCTGGCTTTATTAAGAAATGAAGTTGTTCTGCAATTAATGATACATCAAG 3886
DB 201 CACCTGGCTTTATTAAGAAATGAAGTTGTTCTGCAATTAATGATACATCAAG 260
QY 3887 ATGATTAATCTAGATTTTATTTATTTGCAAAATTAACAAATGTTCTGAGATTAATAAT 3946
DB 261 ATGATTAATCTAGATTTTATTTATTTGCAAAATTAACAAATGTTCTGAGATTAATAAT 320

QY 3947 ACTGCTTATCCAAAGCTAAGTACTAATTTTGGTAAACCAACCACTTTGTTAATATATATG 4006
 DB 321 ACTGCTTATCCAAAGCTAAGTACTAATTTTGGTAAACCAACCACTTTGTTAATATATATG 380
 QY 4007 TAAAGATCATGATTTCCCTTTTGTGTCAGGTTGGGAAAGTTGATGTC 4057
 DB 381 TAAAGATCATGATTTCCCTTTTGTGTCAGGTTGGGAAAGTTGATGTC 431

RESULT 6
 B34439 436 bp DNA linear GSS 17-OCT-1997
 LOCUS HS-1024-B2-B02-MF.abi CIT Human Genomic Sperm Library C Homo
 DEFINITION sapiens genomic clone Plate=CT 803 Col=4 Row=D, genomic survey
 sequence.
 ACCESSION B34439
 VERSION B34439.1 GI:2533808
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 436)
 Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
 Tratsloff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.B.
 Construction of a Characterized Clone Resource for Genomic
 Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
 Tagged Connectors
 Unpublished (1997)
 JOURNAL Contract: Mahairas GG, Zackrone KD, Hood L
 COMMENT University of Washington
 Seattle, WA 98195, USA
 Tel: (206) 616-8744
 Fax: (206) 685-7301
 Email: kzackrone@u.washington.edu
 Sequence Tagged Connector
 Plate: CT 803 row: D column: 4
 Class: BAC ends
 High quality sequence stop: 436.
 Location/Qualifiers
 1. 436
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=CT 803 Col=4 Row=D"
 /sex="M"
 /clone_lib="CIT Human Genomic Sperm Library C"
 /note="Organ: sperm; Vector: pBelBAC11; BAC Clones in
 E-Coli DH10B"

FEATURES
 source

ORIGIN
 Query Match 7.5%; Score 374.8; DB 8; Length 436;
 Best Local Similarity 99.0%; Pred. No. 8.9e-36;
 Matches 387; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 2829 GGAAGCTTGAAGAACTTTGAAGAGTGAAGGCAATTAAGAAAGAAATGCTTTTAC 2888
 DB 47 GCAAGCTTGAAGAACTTTGAAGAGTGAAGGCAATTAAGAAAGAAATGCTTTTAC 106
 QY 2889 CAGGCACTGAATCTTACTTGTGATTAATTTTCTGCTCTTTCTTTTCTCAGCT 2948
 DB 107 CAGGCACTGAATCTTACTTGTGATTAATTTTCTGCTCTTTCTTTTCTCAGCT 166

QY 2949 AACAGACTTAAAGAAAGATCATGTTCCATGATTTTCAACCCAGGGAGCTTCAAGC 3008
 DB 167 AACAGACTTAAAGAAAGATCATGTTCCATGATTTTCAACCCAGGGAGCTTCAAGC 226

QY 3009 TGGCCGTGCCAGAGTTCAAGAGTAAAGCCAGAGTAAAGATGTTTTCAGGA 3068
 DB 227 TGGCCGTGCCAGAGTTCAAGAGTAAAGCCAGAGTAAAGATGTTTTCAGGA 286

QY 3069 AGAGGATCTGAGGACAGGTTTTCAGTGGGTGATGAGCAGCTGATGTTGATCAAGAAGA 3128
 |||||||

DB 287 AAGAGATCTGAGCAGAGGTTTTCAGTGGGTGATGATGACAGCTGATGATCATGAAG-A 345
 QY 3129 TTTAATGAGCTTGTCTACGAGAGCCGCCCTTGTCTTCCAGGCAATTACTGAGCGAGC 3188
 DB 346 TTTAATGAGCTTGTCTACGAGAGCCGCCCTTGTCTTCCAGGCAATTACTGAGCGAGC 405
 QY 3189 CTTCCCAAGTCTGCTGTCATGCTGCTA 3219
 DB 406 CTTCCCAAGTCTGCTGTCATGCTGCTA 436

RESULT 7
 B2879665/ 676 bp DNA linear GSS 18-MAR-2003
 LOCUS CH240_195A22.TJ CHORI-240 Bos taurus genomic clone CH240_195A22,
 DEFINITION genomic survey sequence.
 ACCESSION B2879665
 VERSION B2879665.1 GI:29107067
 KEYWORDS GSS.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovinae; Bos.
 1 (bases 1 to 676)
 Zhao,S., Shetty,U., Shatsman,S., Tsengaye,G., Geer,K.,
 Shvartsbeyn,A., Gebregorgis,E., Chen,D., Riggs,F., de Jong,P.,
 Crawford,A.M. and McEwan,J.C.
 Bovine BAC End Sequences from Library CHORI-240
 Unpublished (2003)
 JOURNAL Contact: Shaying Zhao
 COMMENT Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.choi.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering/information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBMC) by AgResearch Ltd., New Zealand and The
 Institute of Genomic Research (TIGR), USA.
 Plate: 195 row: A column: 22
 Seg primer: SP6
 Class: BAC ends.
 Location/Qualifiers
 1. 676
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="bred: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_195A22"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: PTABAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

FEATURES
 source

ORIGIN
 Query Match 7.2%; Score 359.4; DB 8; Length 676;
 Best Local Similarity 79.6%; Pred. No. 5.3e-34;
 Matches 544; Conservative 0; Mismatches 11; Indels 28; Gaps 9;

QY 2389 GTAGTATCAAAATTTGGGAAATTTCTTGTAAACCAAGGAAATAATCTTGGCTTT 2448
 DB 676 GTAGTATCAAAATTTGGGAAATTTCTTGTAAACCAAGGAAATAATCTTGGCTTT 618

QY 2449 GGGCTGACAGAACTACTTGGCTTGAAGTCAGAGAAAGTCTCTCAAAATCTTAAG 2508
 |||||||

```

Db      617 GGGGTGACGAGAAATTCACCTTCCTTCAGCTCAAGAAAGTATCTCTTGAATCTCTGAA 558
Qy      2509 GTCTTAATATACAGAGCTGAAACTTTAAAGGCAAGCTGCAGTATAGTTGTAATGCTANG 2568
Db      557 GTACGAAATTTAGGAGCTGAACTCAAAAAACAAGCGCGGTAT---TTGGTATCTCATG 501
Qy      2569 GATTGGAACCTTAG--TAATTAAGTCAATATTAATTAAGCAATGACATGA-----TTAAT 2621
Db      500 GATTAAATTTTCAGTTAATTAATTAAGTCAATGAGCTATACAGAAATGCTGACAGAAAGTTAAT 441
Qy      2622 CCCCTACAGCAATTAATTAAGTGAACATGCAAAAAAAGCAAGCACTTAACAGAAAAA 2681
Db      440 CCCCTAAAGAGTAAATTAAGGACAAAGAAAAAAGCTTAGAGTTTAACAGAAAAA- 382
Qy      2682 GTTGCAAAATCATCTCAAGAGATTAGTTAGTTAAGTATGTT- TAAAGCACA- TTTT 2739
Db      381 -TTGGAACCCATCATCAGAGAAATTTAGTTAAGTTAAGTCAATGCTCAAGCAATTTT 323
Qy      2740 AGGTGAAGAAAGATGATGATTTTCAGAGATGATACCATTTAGTCTTTTCAGGGATC 2799
Db      322 AGGTAAAGAAAGATGTTGTTGTTTACAGTATATCATTAAGGTTTTCCTTTT 263
Qy      2800 T-----TTCAAGAAAAGTGCCTTTTGGGGGTACAGAGACTTAAAGAAATTT 2847
Db      262 AAAGAACTGATGTTCAAGAAAGTACATTTTGAAGTGTGGAGAGCTTAAAGACATTT 203
Qy      2848 GAAGAGTGAAGAAAGAGCAATTAAGAAAAATGCTTTTACAGGCACTGATCTTACT 2907
Db      202 GAAGAGTGAAGCAAGAGCAATTAAGAAAAAGGTTTATGCTAGC-CATTTCTTGAAGT 144
Qy      2908 TTGCATAAATTTATTTCTGCTCTTTCTTTTCTCTAGCTAACAGACTTAAAGAGA 2967
Db      143 TTGCATAAATTTATTTCTGCTCTTTCTTTTCTCTAGCTAACAGACTTAAAGAGA 84
Qy      2968 TCATGTTCCATGATCTTCAATTTCAACCCAGAGGACCTTCAAGCTCCCGTCCAGAAATT 3027
Db      83 TCATGTTCCATGATCTTCAATTTCAACCCAGAGGACCTTCAAGCTGATGATGCCAGAAATT 24
Qy      3028 AAGAGTAAAGGCGCAGAGAGTAA 3050
Db      23 AAGAGTAAAGGCGCAGAGAGAA 1

```

```

RESULT 8
LOCUS   CR627133/c 5940 bp mRNA linear HTC 03-AUG-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686J154 (from clone DKFZp686J154).
ACCESSION CR627133
VERSION  CR627133.1 GI:50949747
KEYWORDS HTC.
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 5940)
  Koehler, K., Beyer, A., Mewes, H.W., Weill, B., Amid, C., Osanger, A.,
  Fobo, G., Han, M. and Wiemann, S.
  The German cDNA Consortium
  Direct Submission
  Submitted (13-JUL-2004) MIPS, Ingolstaedter Landstr.1, D-85764
  Neubiberg, GERMANY
  Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
  Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
  sequenced by BMFZ (Biomedical Research Center at the
  Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA
  sequencing consortium of the German Genome Project. This cDNA
  (DKFZp686J154) is available at the RZPD Deutsches Ressourcenzentrum
  fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD
  for ordering:
  http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp686J154
  Further information about the clone and the sequencing project is
  available at http://mips.gsf.de/projects/cdna/.
FEATURES
          Location/Qualifiers

```

```

source      1. 5940
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="DKFZp686J154"
             /csize="type="cervix"
             /clone_id="686 (synonym: h1cc3). Vector pSport1_sfi; host
             DH10B; sites Sfi1A + Sfi1B"
             /dev_stage="adult"
             /note="unclassified"

ORIGIN
Query Match      6.5%; Score 325.2; DB 3; Length 5940;
Best Local Similarity 66.5%; Pred. No. 2.9e-30;
Matches 546; Conservative 0; Mismatches 259; Indels 16; Gaps 5;

Qy      114 GTACAAATATATAAAGATTTTCAGGACGTGGCGTGCACCCAGTCCAGCTATTT 173
Db      4932 GTTAAAGAGTTAAGATATTTGCTGAACACAGGGCTCACCTGTAATCCAGC-ACCT 4874
Qy      174 GGGAGCTGAGGTAGAGGAGATGCTTGAAGCCAGAGTTGAAGACAGCTTAGGCAACT 233
Db      4873 AGGAGCTGAGGTGCTAGAGCCAAATTTAGCCCAAGATTCAAGACAGCTGAGCAACT 4814
Qy      234 AGTGAAGCCGTGCTATTAATA-----AATTAATAGCTGTTGCTTGGACACAGCCCTG 288
Db      4813 GGTAAACCTTCTCTACAAAGATACAAATTAATAGTAGGTGTGTGTGCGCTTG 4754
Qy      289 CAG--CTAGCTCTCGAAGACTGAGTGGAGATCACTGAGCCAGAGGCTAGAGCT 346
Db      4753 TAGTCTTGCTACTGAGGAGCTGAGTGGAGATCACTGAGCCAGAGGCTAGAGCT 4694
Qy      347 GAGAGTGAAGATGATCAACCACTGATTCAGCTCGAAGACAGAGGAGACCTGTTT 406
Db      4693 GAGAGTGAAGATGATGAGATTCACCTCGAGCTGAGCAACAGCAAAACCTGTCA 4634
Qy      407 CCAAAAAAAAAAAAAAAAAAATGCAAGAAAGATCATTAACCTGACCTGGAGACATA 466
Db      4633 CACATGAAGAAAGACATCTGTAATTTAAAGCATCTCTCATTTTCACTCTCTGTA 4574
Qy      467 ACTTTTATGTGATGAATTCACATCTTTTAAAGAAATTTAGCAATTTCTGATAAATGT 526
Db      4573 TATATT-----TATGTTAAATATCTCAAGGAAGAGATTCAAAAACATGAGAAATTA 4520
Qy      527 ATTATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 586
Db      4519 TAAACGACAGAGTAAAGATTGAGAAACAAACATTTGATTTGTTCCATTAATTTGAGACAA 4460
Qy      587 ACTCTCTAGTTGTGAGTCAAAACTTTTATGATGTC--TTTGGCCGGGTGCGGTGCTCAG 644
Db      4459 ACCTCAACTGACTACAAATTTGAGACAAAGAAATCAATTGAGCCGGGCGGGGTCTCAG 4400
Qy      645 CCTGTAAATCCAGCATTTTGGAGGCGGAGCGGGGTGATCAAGGTTTAAAGATTCAG 704
Db      4399 CCTGTAAATCCAGCATTTTGGAGGCGGAGCGGGGTGATGATGATGATGATGATGATGATG 4340
Qy      705 ACCATCTGGGCTAAACAGGTGAACCTGCTCTCTAATAAATAAATAAATAAATAAATAAATA 764
Db      4339 ACCATCTGGGCTAAACAGGTGAACCTGCTCTCTAATAAATAAATAAATAAATAAATAAATA 4280
Qy      765 TGGGTGCGCAGACGCTGTAGTCCAGCTGCTCAGAGAGGCTGAGCAGAGATGATG 824
Db      4279 CGTGTGCGCGGACCTGTATGATCAGAGTACTAGGAGAGCTGAGGACAGAGATGATG 4220
Qy      825 AACCGGAGAGCGAGGCTTGAAGTGAAGCCAGATTTGCGCCACTGCACTCAGCTGGCG 884
Db      4219 AACCTGGAGGAGGAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4160
Qy      885 ACAATGCGAGACTGCTCTCAAAAAAAAAAAAAAAAAAAG 925
Db      4159 ACAAGACGAGACTGCTCTCAAAAAAAAAAATCAATTGAGACAAAG 4119

```

RESULT 9
AO709562/c 472 bp DNA linear GSS 07-JUL-1999
LOCUS HS_5484_B2_B04_T7A RPCI-11 Human Male BAC Library Homo sapiens
DEFINITION genomic clone Plate=1060 Col=8 Row=D, genomic survey sequence.
ACCESSION AO709562
VERSION AO709562.1 GI:5418988
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 472)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
JOURNAL 99380589
MEDLINE 10449764
PUBMED
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering/Bac.htm)
or from Research Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 1060 row: D column: 8
Seq primer: T7
Class: BAC ends
High quality sequence stop: 472.
Location/Qualifiers
1. 472
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=1060 Col=8 Row=D"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"

ORIGIN
Query Match 6.3%; Score 316.2; DB 8; Length 472;
Best Local Similarity 99.1%; Pred. No. 9,6e-29;
Matches 318; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 4422 GGAGTCAGTGTGCATCTCAGATGACGCAACTCTGCTCCTGATCAAGCAATC 4481
DB 359 GAAGTCAGTGTGCATCTCAGATGACGCAACTCTGCTCCTGATCAAGCAATC 300
QY 4482 TCTGTCCTGCTGCTCTTAACTCTTAACTGTTAACTGTTAACTGTTAACTT 4541
DB 299 TCTGTCCTGCTGCTCTTAACTCTTAACTGTTAACTGTTAACTGTTAACTT 240
QY 4542 TTTCATGCTGAAGATCTGCAATGAAGGACCACTTTATTTATTTCAAGCTTC 4601
DB 239 TTTCATGCTGAAGATCTGCAATGAAGGACCACTTTATTTATTTCAAGCTTC 180
QY 4602 ATATCAATTTGGCAGTGAATCTGTTTCAAGGTGGCTCGAA CGGGGCACTGGAAC 4661
DB 179 ATATCAATTTGGCAGTGAATCTGTTTCAAGGTGGCTCGAA CGGGGCACTGGAAC 120

QY 4662 ATACTTGATATGAGGACCAATGACACTTCTGATCTCTTCTGAGTTGACTTGG 4721
DB 119 ATACTTGATATGAGGACCAATGACACTTCTGATCTCTTCTGAGTTGACTTGG 60
QY 4722 ATGTTCTGCACAGACCTTTC 4742
DB 59 ATGTTCTGCACAGACCTTTC 39

RESULT 10
AO839831 6146 bp DNA linear GSS 31-MAR-2000
LOCUS 260L13-C47 C1TB Homo sapiens genomic clone 260L13, genomic survey
DEFINITION sequence.
ACCESSION AO839831
VERSION AO839831.1 GI:6652463
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 6146)
Carpen,J.D., Makolowska,I., Robbins,C.M., Scott,N., Sood,R.,
Comors,T.D., Bonner,T.I., Smith,J.R., Farnue,M.U., Stephan,D.A.,
Pinkert,H., Morgenbesser,S.D., Su,K., Graham,C., Gregory,S.G.,
Williams,H., McDonald,L., Baxevanis,A.D., Klingler,K.W. and
Landes,G.M.
A 6-Mb high-resolution physical and transcription map encompassing
the hereditary prostate cancer 1 (HPC1) region
Genomics 64 (1), 1-14 (2000)
JOURNAL 20175426
MEDLINE 10708513
PUBMED
COMMENT Contact: Carpen JD
Cancer Genetics Branch
National Human Genome Research Institute/National Institutes of
Health
Bldg. 36, Room 3D04, 36 Convent Drive, Bethesda, MD
Tel: 301 435 5626
Fax: 301 435 5465
Email: jdc@hgrl.nih.gov
Class: Shotgun.
Location/Qualifiers
1. 6146
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="260L13"
/clone_lib="C1TB"

ORIGIN
Query Match 6.1%; Score 302.8; DB 8; Length 6146;
Best Local Similarity 63.3%; Pred. No. 1.4e-27;
Matches 579; Conservative 0; Mismatches 297; Indels 38; Gaps 6;
QY 109 TGTCTGTAAATTAATAAAGATTTTCCAGCAGTGTGCGTCAACCCCACTGCAAC 168
DB 943 TATCTAATTAATTTCTAACTACAGAGGCTGGCACCGTCTCATGCTGTAATCCAC 1002
QY 169 TATTTGGAGGCTGAGATGAGAGATGCTTAAGCCAGAGATTGAAGAACAAGCTAGGC 228
DB 1003 ACTTTGGAGGCTGAGATGAGAGATGCTTAAGCTGAGAGATTGAAGAACAAGCTAGGC 1062
QY 229 AACATAGTGAACCTGTGTCTATAAAAT-----AATTAGCTGTGTCTTGGCAG 283
DB 1063 AACATAGTGAACCTGTGTCTATAAAATAGAAAATTAATACCGGCGCTGATGCGACAT 1122
QY 284 GCCTGCA--GTTAGCTACCTGGAAGAAGTGGAGAGATCAC--TGAGCCACAGAGGCT 340
DB 1123 GCCTGTAATCCACGCTACTCAGAGAGCTGAGGACAGAGATCTTTGAACCTGGAGGCA 1182
QY 341 GAGGCTGAGTGAACAGTATCACCCAGCTGATTCAGGCTGGAAGACAGAGAGAGACC 400
DB 1183 GAGGCTGAGTGAACAGTATTCGCAATTCGACCTCGGCTGGTAAATGAGCGAAATC 1242

```

QY 401 CTGTTTCAAAAAAAAAAAAAAAAAATGCAAGAAAGACATCATTAACCTGACCTGG 460
Db 1243 TGCTCCCAAAATATAAATAATATCACTACAGAGATACAGATGATTAATTTTC 1302
QY 461 GACATTACTT-----TATGATGAAATTCACATCTTTTAGAGAAATAG--- 509
Db 1303 TTAACAATTTTGAATATCTATTAATGATTTCTGATCTTTTGAAGAAAGAGA 1362
QY 510 -----CATTTCTGATTAATATGATATTAATATTAATTAATTAATTC 551
Db 1363 GGAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1422
QY 552 AAATGGAATTAATATTTCTGAGAACTAGCTTCTCTCTCTCTCTCTCTCTCTCTCT 611
Db 1423 TGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1482
QY 612 TTAATGATCT-TTGGCGGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 670
Db 1483 AAGTTCAAGTCTTGGCGGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 1542
QY 671 CGAGGCGGGGTGGATCAGAGGTTAGAGATCGAGACCATCTGGCTAAACGCTGAACC 730
Db 1543 CGAGGCGGGGTGGATCAGAGGTTAGAGATCGAGACCATCTGGCTAAACGCTGAACC 1602
QY 731 TCGCTCTACTAATAAATCAAAAAAATTAAGCCGGGTGGGTGGGTGGGTGGGTGGGT 790
Db 1603 CTGCTCTACTAATAAATCAAAAAAATTAAGCCGGGTGGGTGGGTGGGTGGGTGGGT 1662
QY 791 GCTGCTCAGAGGCTGAGGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 850
Db 1663 GCTACTTGGGATGCTGAGGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1722
QY 851 GCCGAGTTTCCGCCACTGCTCACTCCAGCTGGGCGAGATGCGAGACTCTGTCAAAAA 910
Db 1723 GCCAAGATTTGGCCACTGCTCACTCCAGCTGGGCGAGATGCGAGACTCTGTCAAAAA 1782
QY 911 AAAAAAAAAAAAAAGTGAATGCTTTGAGCCAGTCAAGTCTTTCTTTCTTTCTTTCT 970
Db 1783 AAAAAAAAAAAAAAGTGAATGCTTTGAGCCAGTCAAGTCTTTCTTTCTTTCTTTCT 1842
QY 971 TTTTTTTTTTTTTT 984
Db 1843 TTAATAATTTTCTT 1856

RESULT 11
LOCUS AF057104 1300 bp DNA linear GSS 21-FEB-2001
DEFINITION AF057104 Human Homo sapiens genomic clone 51R probe, genomic survey
sequence.
ACCESSION AF057104
VERSION AF057104.1 GI:3142411
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1300)
Tanev,M. and Geraghty,D.E.
The human major histocompatibility complex: 42,221 bp of genomic
sequence, high-density sequence-tagged site map, evolution, and
polymorphism for HLA class I
Genomics 51 (1), 35-44 (1998)
98360090
JOURNAL MEDLINE
PUBMED 98360090
COMMENT Contact: Janer M
Geraghty Lab
Fried Hutchinson Cancer Research Center
1100 Fairview Ave., D2-100, Seattle, WA 98109-1024, USA
Email: geraghty@thrc.org
Class: unknown.
Location/Qualifiers

```

```

source 1..1300
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="6p21.3; HLA Class I region"
/clone="51R probe"
/clone_1ib="Human"

ORIGIN
Query Match 5 9%; Score 293.2; DB 8; Length 1300;
Best Local Similarity 66.2%; Pred. No. 3.7e-26;
Matches 564; Conservative 0; Mismatches 243; Indels 45; Gaps 8;

QY 102 AAACCTCTCTCTGTAACAATAATAAAGATTTTCAGGCGATGATGAGCGGTGACCCCGC 161
Db 38 ATACAAATTTTATTCATAAATAACAAATTAATAGGCTGGCGGGGTGGCTGAGCCTGTRA 97
QY 162 TGGCAGCTATTTTGGAGGCTGAGTGAAGAGATCTTGAAGCCAGAGTTGAAGACAAG 221
Db 98 TCTTAGCATTGGGAGGCGGAGCGGCGAG--TCAGAGGTCAAGAGTTCAAGACCGAG 155
QY 222 CCTAGGCAACATAGTAGAGCCCTGATCTAATAAAAT-----AATTAGCTGTTGCT 275
Db 156 CTTGGCAACATAGTAGAACCCTGATCTACTAATAAATCAAAAAATTAAGCTGGCGTAG 215
QY 276 TGGCAAGGCTGCA--GCTAGCTACTGGAAGACTGAGAGTGGAGATCAC-TGAGCCC 332
Db 216 TGGCGGGGTCTGTATATCCAGCTACTCGGGAGACTAAGGCAACAGATCGCTGAACCC 275
QY 333 AAGAGGCTGAGGCTGCAATGAAACATGATCACCCAGCTGATTCAGCCTGGAGACAGA 392
Db 276 GGAAGCGGAGGTTGACAGTAGACCCAGATCGACATCGACTCCAGCTTGGCCACAGA 335
QY 393 GGAAGACCTGTTTCCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 452
Db 336 GTGAGACTGTGTCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 395
QY 453 TGACCTGGAGACATACTTTATGATGATAATTC--CAATCTTTTGAAGAAATTAAG 509
Db 396 TTTGCCCTTATTAACATTTATTAATATATATTAATTAATTAATTAATTAATTAATTA 455
QY 510 CATTTCTGATTAATAATGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 554
Db 456 CTTAATTTATTTTGAATAATGATGCTTTAACAATTTGTAATGACAGATTGAAGCTCTGG 515
QY 555 -TGAATTAATAATTTCTGAGAAACTAGCTCTCACTCTCAGTTGTCAATCAAACTTT 613
Db 516 TTCTATGTTGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 575
QY 614 AATGATCTTTGGCGGGGTGGCTGCTCAGCCCTGTAATCCAGACATTTGGAGGCCGA 673
Db 576 CTTTGTGATTAACAGAGGCGGCTGCTCAGCCCTGTAATTAACACACTTTTGGAGTCTGA 635
QY 674 GCGGGGTGATCACAAGGTTAGAGATGAGACCATCTGCTAAACCGGTAACCTCG 733
Db 636 GCGGGGCGGATCAGAGATGAGAGATGAGACCATCTGCTAAACCGGTAACCGGTAACCTCG 695
QY 734 TCTTACTAATAAATAAC--AAAAATTAAGCCGGGTGCGGTGTCAGACGCTGTAGTCCAG 791
Db 696 TCTTAATAAATAATACAAAAAATAATTAAGCCAGGTGCGGTGTCAGACGCTGTAGTCCAG 755
QY 792 CTGCTCAGAGAGCTGAGGAGAGAAATGATGTAACCCGAGAGCGGAGCTTGAAGTAG 851
Db 756 CTACTTGAAAGCTGAGGAGAGAAATGATGTAAC-----CTTGAAGTAGAG 802
QY 852 CCGAGATTGCGCCACTGCACTCAGCTTGGGCGAGAGTGGAGAGTCTGTCTCAAAAAA 911
Db 803 CCGAGATTGCGCCACTGCACTCAGCTTGGGCGAGAGTGGAGAGTCTGTCTCAAAAAA 862
QY 912 AAAAAAAAAAAAA 923
Db 863 AAAAAAAAAAAAA 874

```

[illegible]

RESULT 13	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT
BC037921	1279 bp	mRNA	linear	HTC 04-MAR-2003								
BC037921		Homo sapiens, clone IMAGE:5278633, mRNA.										
BC037921	GI:23138806											
HTC.												
Homo sapiens (human)												
Homo sapiens												
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.												
1 (bases 1 to 1279)												
Strausberg, R.												
Direct Submission												
Submitted (16-SEP-2002)												
National Institutes of Health, Mammalian Gene Collection (MGC) Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA												
NIH-MGC Project URL: http://mgc.nci.nih.gov												
Contact: MGC help desk												
Email: cgabs-remail.nih.gov												
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.												
CNS Library Preparation: Michael J. Brownstein (NHGRI) & Shtrakt												
Toshiyuki and Piero Carninci (RIKEN)												
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)												
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada												
info@cgsc.bc.ca												
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Shanna Chan, Readman Chlu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywnski, Reta Kutische, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saeedi, Jacqueline Scheit, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natarja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.												
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov												
Series: IRAX Plate: 75 Row: m Column: 8												
This clone was selected for full length sequencing because it passed the following selection criteria: Excellent library of origin												
This clone has the following problem: retained intron.												
Location/Qualifiers												
1. 1279												
/organism="Homo sapiens"												
/mol_type="mRNA"												
/db_xref="taxon:9606"												
/clone="IMAGE:5278633"												
/tissue_type="Brain, hypothalamus"												
/clone_id="NIH MGC_96"												
/lab_host="DH10B"												
/note="Vector: pBluescript"												
ORIGIN												
Query Match	5.8%	Score 291.6;	DB 3;	Length 1279;								
Best Local Similarity	66.1%	Pred. No. 5.8e-26;										
Matches 522;	Conservative 0;	Mismatches 219;	Indels 49;	Gaps 5;								
139	GGCATGTGTCGCGACCCCGAGGCCAGCATTTTGGAGGCTAGAGGAGGAATGCT	198										
531	GGCAGAGTGGCTACCCCTGTATCCAGTACTTTGGAGGCTAGAGTGGATGATCAC	590										
199	TGAAGCCAGAGTTGAAGACAAGCTTGAAGCAACTAGTGAAGACCCTGTGTATAAAAA	258										

[illegible]

	RESULT	14			
	BMS59057				
LOCUS	BMS59057	647 bp	mRNA	linear	EST-20-FEB-2002
DEFINITION	AGENCOURT_6598315 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5474291				
ACCESSION	5' mRNA sequence.				
VERSION	BMS59057				
KEYWORDS	BMS59057.1 GI:18802313				
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				
TITLE	Homo sapiens				
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheraia; Primates; Carnivora; Homnidae; Hmo.				
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs@rcmail.nih.gov Tissue Procurement: DCTD/DTP CDNA Library Preparation: Rubin Laboratory CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov plate: LNCMI987 row: e column: 12 High quality sequence spot: 603.				

```

FEATURES
source
location/Qualifiers
1..647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon.9606"
/clone="IMAGE:5474291"
/issue_type="amelanotic melanoma, cell line"
/lab_host="DHI0B (pHage-resistant)"
/clone_1b="NIH_MGC_41"
/ncbi="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGGACGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

```

Query Match	5.8%;	Score 288.2;	DB 4;	Length 647;
Best Local Similarity	68.0%;	Pred. No. 2e-25;		
Matches 431; Conservative	0;	Mismatches 199;	Indels 4;	Gaps 2

Qy	292	CTAGCTACTCGAAGACTGAGAGTGGAGATCA-CTGAGCCGAGAGGCTGAGCGTGG	350
Db	8	CCAGCTACTCAGAGAGCGTAAAGCAGAGAGATCACTGAGCTCAGAGAGTGGAGTTAATG	67
Qy	351	TTAAACAGATGATCCCGAGCTGGATTTCCAGCTCTGAAAGACAGAGGAGACCTCTGTTCCAA	410
Db	68	TTAGCTGAGATTTGTGCCATTTGACATCCAGCTTGGGTGACAGAGTGTGACCTGTCTCAAA	127
Qy	411	AAAAAAAAAAAAAAAAATGCAAGAAAACATCACTAACTTGACCTGGACATTACTT	470
Db	128	AAAAATTTTTTTTTTAAACAGTTGACAGAGGATTTAAAGATTTTTTTCATTTGAT	187
Qy	471	TTATGTGATGAAATTCACAACTTTTAGAAGAAATTAGCATTTCTGATTTAAATGTATTA	530
Db	188	CAAGTTGAAAACCAAGGCAAGGCTTAGCATTTTTTTTAGGTGACACGATPAAACAGTGA	247
Qy	531	TAATTATTTATTTAAATTCMAATGGAATTAATATTTCTGAGAACTAGCTTCTCACTC	590
Db	248	GAAAGTATGAGGACTATATGACAAATTAAGTTTGGGGTGGCGAGCATTCATGTAGGAT	307
Qy	591	TCTCGTTGTGAGCAAAACCTTTAATGCTTTTGGCGGGTGGGGTCAAGCGCTGA	650
Db	308	GGGACACAAACCTTGAAGTTAGTG--TAGGGCGGGCGGGTGGCTCAGCGCTGTA	364
Qy	651	ATCCAGCATTTTGGAGGCCGAGGCGAGTGGATCAAGATTTAGAGATGAGACATC	710
Db	365	ATCCAGCATTTTGGAGGCCGAGGCGGGTGAATCATGAGGTACAGACATCAGACCAAC	424
Qy	711	CTGGCTTAAACAGGTGAACCTCGTCTCTACTATAAAATACAAAAATTAGCGGGGCGGT	770
Db	425	CTGGCTTAAACAGGTGAACCTCGTCTCTACTATAAAATACAAAAATTAGCGGGGCGGT	484
Qy	771	GCACAGCGCTGTAGTCTCCAGCTGCTCTAGAGAGCTGAGGACAGAGAAATGTTGTGAACCG	830
Db	485	GGCGGGCGCCGTGATGCTCCAGCTTACTCTGGAGGCTGAGGACAGAGAAATGCTGTGAACCG	544
Qy	831	GGAGCGGAGCTTCACTGAGCGGAGATTTGCGCACTGCACTTCACAGCTGGGGCGACAGT	890
Db	545	GGAGCGGAGCTTCACTGAGCGGAGATTTGCGCACTGCACTTCAGAGCTGGGGCGACAGAG	604
Qy	891	CGAGACTCTGCTCAAAAAAAAAAAAAAAAAAAAAA	924
Db	605	CGAGACTCTGCTCAAAAAAAAAAAAAAAAAAAAAA	638

RESULT 15	AL120269/c	660 bp	mRNA	linear	EST 04-SEP-2003
LOCUS	AL120269				
DEFINITION	DKFZp761K027.g1.761 (synonym: hamy2) Homo sapiens cDNA clone DKFZp761K027.3, mRNA sequence.				

ACCESSION	AL120269
VERSION	AL120269.1
KEYWORDS	GI:5926168
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 660)
JOURNAL	Koehter K., Beyer A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
COMMENT	EST (Koehter, et al.)
	Unpublished (1999)
	Contact: MIPS

ORIGIN

Query Match	5.7%;	Score 284.8;	DB 1;	Length 660;
Best Local Similarity	69.7%;	Pred. No. 5e-25;		
Matches 463; Conservative	0;	Mismatches 182;	Indels 19;	Gaps 5;

QY	262	TTAGCTGGTTGCTCTTGCGACAGAGCCCTGC/-	-GCTAGCTACTCGGAAGCTGAGAGTGGAG	319
Db	648	TAAGACGACGTGTGCACACA	TGCTCTGTAAACCCCACTATTGGAGGCTGAGGTGGAG	589
QY	320	GATCAC-TGAGCCCGAGAGCGCTGAGGCTGCAGTGAACAGTGATCACCAGCTGATTCGA		378
Db	588	GATCGCTTGAAGCCCGAGAGTTTGAAGGCTCGGCTGAGCTATAATTATGCACTGCATCCCA		529
QY	379	GCCCTGGAAGACAGAGGAGACCCCTGTTCCAAAAA	AAAAAAAAAAAAAAAAATGCAGAAA	438
Db	528	GCCCTGGGACACAGTAAAGACCCCACTTC/-	TTAAAAAGTTAAAAATAAATAAAGAAAT	471
QY	439	AGACATCATAAACCTTGACCTGGGACATAA	CTTTATGTGATGAAATTCACAACTCTTTAG	498
Db	470	AATCCACTCMAAGCTTAAATGAAAAAGTGATATA	----AAATCATATGCTATTTTAGGT	415
QY	499	GAAGAAATTAGCAATTTCTGATAAAAATGATTAATTAATATATATTTAAATTCAAATGGA		558
Db	414	TATTTTTTCATTAATCTTATTATTAAACATGTAAC	CTTCAGTTGATATGATCTGTATATATA	355
QY	559	ATTAAATTTCTGAGAAACTAGCTCTCACTCTC	CAGTTTCAGTCAACAAACCTTTAATGG	618
Db	354	ATCTATAAACCATAATATATGTAATACAGCTTTGACATTTGAAACTAT	-----T	305
QY	619	TCTTTGGCCGGGTGCGGTGGCTCAAGCCCTGATATCC	ACACACTTTGGAGGCCGAGGCG	678
Db	304	TACTAGGCTGGGCGGGCTGGCTCAAGCCCTGATATCC	CACACACTTTGGAGGCCGAGGTGG	245
QY	679	GTGATACCAAGTTTAGAGATTCGAGACCATCTG	GCTTAACACGCTGAAACCTCGTCTCT	738
Db	244	GCGGATACGAGGTATAGAGAGATCGAGATATCTCTG	CTTAACACAGTAAATCTGCATCTCT	185

QY	73	ACATATAATACAAAAAATTAGCCGGGTGGGGTCCAGACGGCTGTATAGCCAGCTGTCTCA	798
Db	184	ACTATAAATAATCAAAAAATTGGCCAGGCTGGTGTGGGTGCTGTAGTCCAGCTACTTG	1233
QY	799	GGAGGCTGAGCAGAGAAATGGTGTGAACCCGGAGGCGGAGCTTGCAGTGAGCCGAGAT	858
Db	124	GGAGGCTGTGGCAGGAGAAATGGCGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCTGAGAT	65
QY	859	TGGGCCATTCGCACTCCAGGCTTGGGGGACACAGTGCAGACTCTGTCTCAAAAAAAAAAAAAA	918
Db	64	CGGTCCACTCGCACTCCAGGCTTGGGGGACAGAGTGAATTTGTCTAAAAAAAAAAAAAAAAA	5
QY	919	AAAA 922	
Db	4	AAAA 1	

Search completed: December 27, 2004, 10:56:44
Job time : 9570.51 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2004, 17:41:01 ; Search time 256.615 Seconds
(without alignments)
13849.312 Million cell updates/sec

Title: US-10-786-065-3_COPY_1_5000

Perfect score: 5000
Sequence: 1 tataggccaatgcgtgctgctc.....ctatcttcttatctgct 5000

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/pdata/1/ina/5A_COMB.seq: *
2: /cgn2_6/pdata/1/ina/5B_COMB.seq: *
3: /cgn2_6/pdata/1/ina/6A_COMB.seq: *
4: /cgn2_6/pdata/1/ina/6B_COMB.seq: *
5: /cgn2_6/pdata/1/ina/PCUS_COMB.seq: *
6: /cgn2_6/pdata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5000	100.0	53332	4	US-09-801-861-3 Sequence 3, Appl1
2	5000	100.0	53332	4	US-10-224-562-3 Sequence 3, Appl1
3	336.2	6.7	17000	4	US-09-679-299A-18 Sequence 18, Appl1
4	334.4	6.7	84495	3	US-09-797-906-3 Sequence 3, Appl1
5	333.8	6.7	45546	3	US-09-146-053-6 Sequence 6, Appl1
6	319.2	6.4	16063	3	US-09-801-052-3 Sequence 3, Appl1
7	319.2	6.4	16063	4	US-10-020-121-3 Sequence 3, Appl1
8	317	6.3	107820	4	US-09-792-616-1 Sequence 1, Appl1
9	309	6.2	30350	4	US-10-118-328-3 Sequence 3, Appl1
10	303.6	6.1	15297	3	US-09-817-180-3 Sequence 3, Appl1
11	303.6	6.1	15297	4	US-10-003-295-3 Sequence 3, Appl1
12	292.4	5.8	63000	4	US-09-759-359A-18 Sequence 18, Appl1
13	292	5.8	70000	4	US-09-851-896-3 Sequence 3, Appl1
14	287	5.7	66933	4	US-09-544-398B-11 Sequence 11, Appl1
15	287	5.7	66933	4	US-09-543-771-11 Sequence 11, Appl1
16	285.4	5.7	72049	4	US-09-544-398B-9 Sequence 9, Appl1
17	285.4	5.7	72049	4	US-09-543-771-9 Sequence 9, Appl1
18	283.8	5.7	26928	4	US-09-544-398B-6 Sequence 6, Appl1
19	283.8	5.7	26928	4	US-09-543-771-6 Sequence 6, Appl1
20	282.4	5.6	114793	4	US-10-148-806-3 Sequence 3, Appl1
21	281.8	5.6	1000	3	US-09-018-584A-33 Sequence 33, Appl1
22	281.8	5.6	1000	4	US-09-784-423-33 Sequence 33, Appl1
23	276	5.5	81001	4	US-09-750-580-1 Sequence 1, Appl1
24	274.4	5.5	90541	4	US-09-759-359A-3 Sequence 3, Appl1
25	274.4	5.5	90541	4	US-10-207-973-3 Sequence 3, Appl1
26	274.2	5.5	114793	4	US-10-148-806-3 Sequence 3, Appl1
27	274.2	5.5	162450	3	US-09-345-882-1 Sequence 1, Appl1

28	272.6	5.5	161652	4	US-09-497-855A-40 Sequence 40, Appl1
29	272	5.4	51719	4	US-09-918-686-2 Sequence 2, Appl1
30	272	5.4	92139	4	US-09-918-686-1 Sequence 1, Appl1
31	271.4	5.4	12619	4	US-09-616-289-49 Sequence 49, Appl1
32	270.4	5.4	8220	4	US-09-797-908-3 Sequence 3, Appl1
33	270.4	5.4	43950	3	US-09-735-934A-3 Sequence 3, Appl1
34	270.4	5.4	43950	4	US-10-060-332-3 Sequence 3, Appl1
35	270	5.4	43950	4	US-10-339-657-3 Sequence 3, Appl1
36	270	5.4	193303	4	US-09-497-855A-37 Sequence 37, Appl1
37	270	5.4	193303	4	US-09-497-855A-44 Sequence 44, Appl1
38	269.8	5.4	90541	4	US-09-759-359A-3 Sequence 3, Appl1
39	269.8	5.4	90541	4	US-10-207-973-3 Sequence 3, Appl1
40	269.6	5.4	18000	4	US-09-657-346A-17 Sequence 17, Appl1
41	268.2	5.4	22481	3	US-08-367-841A-43 Sequence 43, Appl1
42	268.2	5.4	22481	5	PCT-US95-07201-43 Sequence 43, Appl1
43	268.2	5.4	22481	5	US-09-875-223-2 Sequence 2, Appl1
44	268.2	5.4	22484	4	US-09-875-114-2 Sequence 2, Appl1
45	268.2	5.4	254366	4	US-09-822-871-3 Sequence 3, Appl1

ALIGNMENTS

RESULT 1									
US-09-801-861-3									
Sequence 3, Application US/09801861									
Patent No. 6492154									
GENERAL INFORMATION:									
APPLICANT: YAN, Chunhua et al.									
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC									
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES									
FILE REFERENCE: CLO01098									
CURRENT APPLICATION NUMBER: US/09/801,861									
CURRENT FILING DATE: 2001-03-09									
NUMBER OF SEQ ID NOS: 10									
SOFTWARE: FastSeq for Windows Version 4.0									
SEQ ID NO 3									
LENGTH: 53332									
TYPE: DNA									
ORGANISM: Human									
US-09-801-861-3									
Query Match 100.0%; Score 5000; DB 4; Length 53332;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	TATAGGCCAATGCTGCTCAGCGCTGATATCCAGCACTTGGAGGACGAGATCG	60						
DB	1	TATAGGCCAATGCTGCTCAGCGCTGATATCCAGCACTTGGAGGACGAGATCG	60						
QY	61	CTTGAAGCTGAGAAATGGAGCAACCTGTAATAGTGAACCTCTGTGTCAAA	120						
DB	61	CTTGAAGCTGAGAAATGGAGCAACCTGTAATAGTGAACCTCTGTGTCAAA	120						
QY	121	TATATAAGAAATTTTCAGGCAATGCTGCTGACCCCACTGCACTATTGGAGGC	180						
DB	121	TATATAAGAAATTTTCAGGCAATGCTGCTGACCCCACTGCACTATTGGAGGC	180						
QY	181	TAGATAGAGAGAAATGCTTGAAGCAGAGCTGAAGCAAGCTGGCAACTAGTGA	240						
DB	181	TAGATAGAGAGAAATGCTTGAAGCAGAGCTGAAGCAAGCTGGCAACTAGTGA	240						
QY	241	CCCTGTGCTATATAAATTAATTAAGCTGTTCTTGGACAGGCTGAGCTAGTACT	300						
DB	241	CCCTGTGCTATATAAATTAATTAAGCTGTTCTTGGACAGGCTGAGCTAGTACT	300						
QY	301	CGGAAGACTGAGCTGAGAGATCACTGAGCCAGAGAGGCTGAGTGAACGTGA	360						
DB	301	CGGAAGACTGAGCTGAGAGATCACTGAGCCAGAGAGGCTGAGTGAACGTGA	360						
QY	361	TCACCAGCTGATTCAGGCTGGAAGACAGAGGAGACCTGTTCCAAAAA	420						
DB	361	TCACCAGCTGATTCAGGCTGGAAGACAGAGGAGACCTGTTCCAAAAA	420						

Db 361 TCACCAAGCTGATTCAGCCTGAGAGACAGAGGAGACCCCTGTTTCCAAAAA 420
QY 421 AAAAAAAAAATGCAAGAAAAGACATCATTAACCTTGACATTAATTTATGTGATG 480
Db 421 AAAAAAAAAATGCAAGAAAAGACATCATTAACCTTGACATTAATTTATGTGATG 480
QY 481 AAAAAAAAAATGCAAGAAAAGACATCATTAACCTTGACATTAATTTATGTGATG 540
Db 481 AAAAAAAAAATGCAAGAAAAGACATCATTAACCTTGACATTAATTTATGTGATG 540
QY 541 AAAAAAAAAATGCAAGAAAAGACATCATTAACCTTGACATTAATTTATGTGATG 600
Db 541 AAAAAAAAAATGCAAGAAAAGACATCATTAACCTTGACATTAATTTATGTGATG 600
QY 601 CAGTCAAACTTTAATGCTTTTGGCCGGGTGCGTGGCTCAAGCTGTAATCCAGCAC 660
Db 601 CAGTCAAACTTTAATGCTTTTGGCCGGGTGCGTGGCTCAAGCTGTAATCCAGCAC 660
QY 661 TTTGGGAGCGGAGGCGGAGTGCACAGAGTTAGAGATCGAGACATCCCTGGCTAAC 720
Db 661 TTTGGGAGCGGAGGCGGAGTGCACAGAGTTAGAGATCGAGACATCCCTGGCTAAC 720
QY 721 CCGTGAAACCTCGCTCTACTAATAAATAACAAAAATTTAGCCGGGTGGTGCGAGCC 780
Db 721 CCGTGAAACCTCGCTCTACTAATAAATAACAAAAATTTAGCCGGGTGGTGCGAGCC 780
QY 781 TGTATGCCAGCTGCTCAGAGGCTGAGGCGAGAGATGCTGTGAACCCGGAGGCGAG 840
Db 781 TGTATGCCAGCTGCTCAGAGGCTGAGGCGAGAGATGCTGTGAACCCGGAGGCGAG 840
QY 841 CTGGAGTGGCGGAGATTTGGCGCATCGCATCTCAGCTGGGGGAGAGAGCGAGACTCG 900
Db 841 CTGGAGTGGCGGAGATTTGGCGCATCGCATCTCAGCTGGGGGAGAGAGCGAGACTCG 900
QY 901 TCTCAAAAAAAGTGAATGCTTGTGAGCCAGTAAGCTTCTCTTTC 960
Db 901 TCTCAAAAAAAGTGAATGCTTGTGAGCCAGTAAGCTTCTCTTTC 960
QY 961 TCTCTCTCTCTTTTCTTTTCTTCAAAAAATCTCTAGATTGAATCTTGGAAATGGCT 1020
Db 961 TCTCTCTCTCTTTTCTTTTCTTCAAAAAATCTCTAGATTGAATCTTGGAAATGGCT 1020
QY 1021 TAAAGTCTCTCTCTTGTGGCAATTTTGAATGAATAATCATGCTCATTAATAATTAC 1080
Db 1021 TAAAGTCTCTCTCTTGTGGCAATTTTGAATGAATAATCATGCTCATTAATAATTAC 1080
QY 1081 CTGAACATTTTAAAAAACAATCATGAGGTTCAAAATATCAAAATATCATTAATTTGTGT 1140
Db 1081 CTGAACATTTTAAAAAACAATCATGAGGTTCAAAATATCAAAATATCATTAATTTGTGT 1140
QY 1141 GATTAATGACATTAATTTTCTTAAATTAATGATTTGTTATATATCTCCATTC 1200
Db 1141 GATTAATGACATTAATTTTCTTAAATTAATGATTTGTTATATATCTCCATTC 1200
QY 1201 TGTCTCACTTATGATTAATTAATTAATGAGCAATCTTGAAGATCTTAACAGAGAA 1260
Db 1201 TGTCTCACTTATGATTAATTAATTAATGAGCAATCTTGAAGATCTTAACAGAGAA 1260
QY 1261 AGTGTTCATTTAGAGACATGACATGCAACCAATTTAAGCCAGATTCCTTGACATGTT 1320
Db 1261 AGTGTTCATTTAGAGACATGACATGCAACCAATTTAAGCCAGATTCCTTGACATGTT 1320
QY 1321 GTGCTGTAAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1380
Db 1321 GTGCTGTAAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1380
QY 1381 CAAAGCAAGAAAGAAAGAAATTTTCAATCAAACTCTCTCCATGACCATGGCTA 1440
Db 1381 CAAAGCAAGAAAGAAAGAAATTTTCAATCAAACTCTCTCCATGACCATGGCTA 1440
QY 1441 AATATCATATTTGTACAGTTAAGAACATATAGTGTCTCACCAATATGTTTTGAATAA 1500
Db 1441 AATATCATATTTGTACAGTTAAGAACATATAGTGTCTCACCAATATGTTTTGAATAA 1500

QY 1501 TGAATGAATGGCAACCTTCTAAGACTATTTGATACATATTTGTTGAAGGCAAGAGAT 1560
Db 1501 TGAATGAATGGCAACCTTCTAAGACTATTTGATACATATTTGTTGAAGGCAAGAGAT 1560
QY 1561 GCAATGATATTTTCAACTTTTCTGTTTATATTTCTGAGTTCTTGAATCTAA 1620
Db 1561 GCAATGATATTTTCAACTTTTCTGTTTATATTTCTGAGTTCTTGAATCTAA 1620
QY 1621 AAGTTAGTAGGTAAGCAATTTGTTTAAAGTCTGAAAACCAAAATGCTTCAAGATAA 1680
Db 1621 AAGTTAGTAGGTAAGCAATTTGTTTAAAGTCTGAAAACCAAAATGCTTCAAGATAA 1680
QY 1681 GGTAGGAGAAAAATCTCTCAACATGTCATTTAGACCAAGAAAACCTAATATCAA 1740
Db 1681 GGTAGGAGAAAAATCTCTCAACATGTCATTTAGACCAAGAAAACCTAATATCAA 1740
QY 1741 TATCACCATCAATGATATCAATTAATATCAATTTGATAGTAAGCAATGCAATCCCTAA 1800
Db 1741 TATCACCATCAATGATATCAATTAATATCAATTTGATAGTAAGCAATGCAATCCCTAA 1800
QY 1801 AAATATGATATCAATATAGCACTAACTTGTGCGCAAGAACAAACCTTAATCTGCGAAA 1860
Db 1801 AAATATGATATCAATATAGCACTAACTTGTGCGCAAGAACAAACCTTAATCTGCGAAA 1860
QY 1861 TTTTATCTTATCAATTAACAGCTGCTGTTTCAAGTTGTGCAATCTGAATGCAAGCAA 1920
Db 1861 TTTTATCTTATCAATTAACAGCTGCTGTTTCAAGTTGTGCAATCTGAATGCAAGCAA 1920
QY 1921 TCCCTGTCTGATGATGATTTCTTCACTGATTAAGAAAATGCTGAAATGTTAGAGGT 1980
Db 1921 TCCCTGTCTGATGATGATTTCTTCACTGATTAAGAAAATGCTGAAATGTTAGAGGT 1980
QY 1981 GCTCAGAGCAAGCCATCATGATGATATGAAAGCTTCAAGCTGATGCTTGGCAA 2040
Db 1981 GCTCAGAGCAAGCCATCATGATGATATGAAAGCTTCAAGCTGATGCTTGGCAA 2040
QY 2041 AAGGAGAGAAACAAGATTAAGAGAGCAATGGGGGAAAGTTCAAGTGGGGTTTCTC 2100
Db 2041 AAGGAGAGAAACAAGATTAAGAGAGCAATGGGGGAAAGTTCAAGTGGGGTTTCTC 2100
QY 2101 CTGGAACCTAACAAGATTAAGAGCTGTGTGCAAAAGCTGATACAGCTGATGCA 2160
Db 2101 CTGGAACCTAACAAGATTAAGAGCTGTGTGCAAAAGCTGATACAGCTGATGCA 2160
QY 2161 GCTGTACCATTTGTTCAAGAGGAGGCGAGGCAACAAGCTGTGAGAGTCAAGCTGAGA 2220
Db 2161 GCTGTACCATTTGTTCAAGAGGAGGCGAGGCAACAAGCTGTGAGAGTCAAGCTGAGA 2220
QY 2221 CCAAGGAGGCTGCTTCAAGGATTTGATCCCAAGAGAGACATCTTCTATTTGATA 2280
Db 2221 CCAAGGAGGCTGCTTCAAGGATTTGATCCCAAGAGAGACATCTTCTATTTGATA 2280
QY 2281 AACCAAGAGTTCAAGACATCCCTTTTGTAGCGGATCTGATCTTCTGCGGTAGGCT 2340
Db 2281 AACCAAGAGTTCAAGACATCCCTTTTGTAGCGGATCTGATCTTCTGCGGTAGGCT 2340
QY 2341 AAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2400
Db 2341 AAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2400
QY 2401 TTTGGGGAATTTCTTGTAAACAAGAAAGGAAAAAATTCCTTGGCTTGGCTGACAAA 2460
Db 2401 TTTGGGGAATTTCTTGTAAACAAGAAAGGAAAAAATTCCTTGGCTTGGCTGACAAA 2460
QY 2461 ACTCACTTGGCTTGAAGTCAAGAAAGTGTCTCTCAAAATCTCTAAGGTCTTAATTAAC 2520
Db 2461 ACTCACTTGGCTTGAAGTCAAGAAAGTGTCTCTCAAAATCTCTAAGGTCTTAATTAAC 2520
QY 2521 AGAGCTGAAACTTAAGGCAAGCTGACAGTATTAAGTTGGTATGCTATGAAATTTGAACTT 2580
Db 2521 AGAGCTGAAACTTAAGGCAAGCTGACAGTATTAAGTTGGTATGCTATGAAATTTGAACTT 2580

Qy	2561	TAGTAAATTAGTTCATGATTAATTAGCAATGCCATAGATTAATTCCTCAACAGCAATTAATTA	2640
Db	2561	TAGTAAATTAGTTCATGATTAATTAGCAATGCCATAGATTAATTCCTCAACAGCAATTAATTA	2640
Qy	2641	AGTGGACATGAAAAAAGAACCCAGACTTAACAGAAAAAGTTGCAAAAATCATCATCA	2700
Db	2641	AGTGGACATGAAAAAAGAACCCAGACTTAACAGAAAAAGTTGCAAAAATCATCATCA	2700
Qy	2701	AGAGATTTAGGTTAACCTGAAATGTTAAAGACATTTTTTAGGTGAAGAAAGATGTAGTA	2760
Db	2701	AGAGATTTAGGTTAACCTGAAATGTTAAAGACATTTTTTAGGTGAAGAAAGATGTAGTA	2760
Qy	2761	TTTTCAAGAGTTGATTAACCTTAATGCTTTTTCAAGGATCTTTCAAGAAAATGCTTTTTG	2820
Db	2761	TTTTCAAGAGTTGATTAACCTTAATGCTTTTTCAAGGATCTTTCAAGAAAATGCTTTTTG	2820
Qy	2821	GGGGTACAGAAAGCTTAGAAAAAATTGAGAGTGAATAAGGCAAAATAAGAAAAAT	2880
Db	2821	GGGGTACAGAAAGCTTAGAAAAAATTGAGAGTGAATAAGGCAAAATAAGAAAAAT	2880
Qy	2881	GGTTTTACAGAGCACTGATCTTTACCTTGCATAAATTTATTTCTGCTCTTTCTTTTT	2940
Db	2881	GGTTTTACAGAGCACTGATCTTTACCTTGCATAAATTTATTTCTGCTCTTTCTTTTT	2940
Qy	2941	CTTAAGCTTAACAGACTTAAGAAAGACATCATGTTCCATGACTTCATTTCACCCCAGGGGA	3000
Db	2941	CTTAAGCTTAACAGACTTAAGAAAGACATCATGTTCCATGACTTCATTTCACCCCAGGGGA	3000
Qy	3001	CTTCAAGCTGCCCGTCCCAAGTCTGAAGTGAAGTAAAGGCCACGAGTAACAGATTTGT	3060
Db	3001	CTTCAAGCTGCCCGTCCCAAGTCTGAAGTGAAGTAAAGGCCACGAGTAACAGATTTGT	3060
Qy	3061	TTTTCAAGAAAGAGATCTGAGGCAAGGTTTTCACTGGGTGAATGAGCAGACTGATGTTGATC	3120
Db	3061	TTTTCAAGAAAGAGATCTGAGGCAAGGTTTTCACTGGGTGAATGAGCAGACTGATGTTGATC	3120
Qy	3121	AAGAAAGATTTAATGTGAGCTTTGCTCAACGAGCCGCGCTTGCTCCAGGGCAATTAATCT	3180
Db	3121	AAGAAAGATTTAATGTGAGCTTTGCTCAACGAGCCGCGCTTGCTCCAGGGCAATTAATCT	3180
Qy	3181	GAGGAGGCTTCCCAAGTCTGCTGCGGAATGCTGTCTTAATTTCCCTGGGGAATAAAAGT	3240
Db	3181	GAGGAGGCTTCCCAAGTCTGCTGCGGAATGCTGTCTTAATTTCCCTGGGGAATAAAAGT	3240
Qy	3241	CAACACTTAAAAAAGTGTCTTTCTCTCTCCCTTCAACCGGCTCTTTTCCCATCTCC	3300
Db	3241	CAACACTTAAAAAAGTGTCTTTCTCTCTCTCCCTTCAACCGGCTCTTTTCCCATCTCC	3300
Qy	3301	CTTAAGCAGAGAAAGAGCTTCCCTTTTGGGGCAGGCTCATCTTAATTGAATCTTGAGAA	3360
Db	3301	CTTAAGCAGAGAAAGAGCTTCCCTTTTGGGGCAGGCTCATCTTAATTGAATCTTGAGAA	3360
Qy	3361	GCTGGGTGCAAGGCTCTTAATGCGACAGTTTACAAGGGGATTAGCAGGTGAAGTGAACATAG	3420
Db	3361	GCTGGGTGCAAGGCTCTTAATGCGACAGTTTACAAGGGGATTAGCAGGTGAAGTGAACATAG	3420
Qy	3421	CTGGGAAGAGCTTTAGAGATGAGAGTCCCGGCGCCGCCCAATTTCAATTAATAAGCAGGT	3480
Db	3421	CTGGGAAGAGCTTTAGAGATGAGAGTCCCGGCGCCGCCCAATTTCAATTAATAAGCAGGT	3480
Qy	3481	GAGAATCATTAAGAAAGTTCAATGACACTCAGACCTGTGCAAGACACATAGCCGACAGGGA	3540
Db	3481	GAGAATCATTAAGAAAGTTCAATGACACTCAGACCTGTGCAAGACACATAGCCGACAGGGA	3540
Qy	3541	GAGAGACATGTAATCTTAACAAGCTTGAAGAAAAAACAACCTGCGCTGCTTAATTTAA	3600
Db	3541	GAGAGACATGTAATCTTAACAAGCTTGAAGAAAAAACAACCTGCGCTGCTTAATTTAA	3600
Qy	3601	AATCAGGCCCATTTAAATGTTATATCAGGCTTTCCCTTCTTGCAATTCATTCAGAGAAATTTCA	3660
Db	3601	AATCAGGCCCATTTAAATGTTATATCAGGCTTTCCCTTCTTGCAATTCATTCAGAGAAATTTCA	3660
Qy	3661	AAGAAATTAAGACATTTCTCTACTAGACCAAAAGAACATTAATGACATCTTCAGGCGTGTG	3720

Db	3661	AAGAAATAGACATTCTCTACTGACCCAAAGAAATATATCACTTTCAGGCTGTG	3720
Qy	3721	GGAGGCAACAGTGGTAAAGCGCTCTCAACAGTTTTTTATATCCCTCCATAACAAAT	3780
Db	3721	GGAGGCAACAGTGGTAAAGCGCTCTCAACAGTTTTTTATATCCCTCCATAACAAAT	3789
Qy	3781	GACAGAGTTTGTATAAGCAACCTGGAAATTTGCTGCTTCAATCTCCACCTGGCTTAT	3840
Db	3781	GACAGAGTTTGTATAGCAACCTGGAAATTTGCTGCTTCAATCTCCACCTGGCTTAT	3840
Qy	3841	AGAGAAACGTAAGTGGTTCTGCAAAATTAATGCTACATGCAAAAGATGATAATCTGAG	3900
Db	3841	AGAGAAACGTAAGTGGTTCTGCAAAATTAATGCTACATGCAAAAGATGATAATCTGAG	3900
Qy	3901	ATTTTTTATATTTGCCAAAATACACAAAATGCTGAGAAATTAATAATCTGCTTACAA	3960
Db	3901	ATTTTTTATATTTGCCAAAATACACAAAATGCTGAGAAATTAATAATCTGCTTACAA	3960
Qy	3961	AGCTAAGTACTAATTTTGGTAAACACCAACTTTGTTAAATATGTAAAAAGATCATGA	4020
Db	3961	AGCTAAGTACTAATTTTGGTAAACACCAACTTTGTTAAATATGTAAAAAGATCATGA	4020
Qy	4021	ATTCCTCTTTTAGTCAAGTGGGAAAGTGGATGGCTTTTCTTTATAGTTACATCCA	4080
Db	4021	ATTCCTCTTTTAGTCAAGTGGGAAAGTGGATGGCTTTTCTTTATAGTTACATCCA	4080
Qy	4081	ATAGAGAAAGATATAGGCTCAATAGTGGTAAATATTAATTTAAATAATATAGCTGAT	4140
Db	4081	ATAGAGAAAGATATAGGCTCAATAGTGGTAAATATTAATTTAAATAATATAGCTGAT	4140
Qy	4141	CCGAGTCAAGTGGTGTATTAACAATACTAGTATCAACAACAGTTACAGATTTCTTGTTCCT	4200
Db	4141	CCGAGTCAAGTGGTGTATTAACAATACTAGTATCAACAACAGTTACAGATTTCTTGTTCCT	4200
Qy	4201	TCTCCACTCCCACTGCTTCACTTAACCTGGCCAAAAGAAAAAATTTATATATA	4260
Db	4201	TCTCCACTCCCACTGCTTCACTTAACCTGGCCAAAAGAAAAAATTTATATATA	4260
Qy	4261	CTACTACAAGCTAAATATTTATATTTATCTTAGTATTTATGCTGTATTAATTTAT	4320
Db	4261	CTACTACAAGCTAAATATTTATATTTATCTTAGTATTTATGCTGTATTAATTTAT	4320
Qy	4321	ACTTGTAAAAACAGATTTGTAAGGGACATAACAGTTTTATTTATTTATATATATAT	4380
Db	4321	ACTTGTAAAAACAGATTTGTAAGGGACATAACAGTTTTATTTATTTATATATATAT	4380
Qy	4381	TTATTTATTTATTTTGGATGGAATCTCTGTCAACCCAGCTGAGTCAAGTGGCGATC	4440
Db	4381	TTATTTATTTATTTTGGATGGAATCTCTGTCAACCCAGCTGAGTCAAGTGGCGATC	4440
Qy	4441	TGCAATACGCAACCTGCGCTCTGAGTTCAAGCAACTCTCGCCGCCGCGCTTTA	4500
Db	4441	TGCAATACGCAACCTGCGCTCTGAGTTCAAGCAACTCTCGCCGCCGCGCTTTA	4500
Qy	4501	TACTTTCTTAATCTGTTTAGTCAATGCTGTAACCTTAATTTTCAATGCTGAGAACATC	4560
Db	4501	TACTTTCTTAATCTGTTTAGTCAATGCTGTAACCTTAATTTTCAATGCTGAGAACATC	4560
Qy	4561	TGCAATAAAGGACCAACTTTTATTTATTTCAAGCTTCTCATATCAATTTGGCCATGCT	4620
Db	4561	TGCAATAAAGGACCAACTTTTATTTATTTCAAGCTTCTCATATCAATTTGGCCATGCT	4620
Qy	4621	AACTGTTTCAAGGTGCTGGAACGGGGGACCTGTGAACATACTTGATACATGGGCA	4680
Db	4621	AACTGTTTCAAGGTGCTGGAACGGGGGACCTGTGAACATACTTGATACATGGGCA	4680
Qy	4681	CCATGACACATTGATCTCTCTCTGAGTTCTGAATTTGATTTCTGACACAGACTT	4740
Db	4681	CCATGACACATTGATCTCTCTCTGAGTTCTGAATTTGATTTCTGACACAGACTT	4740
Qy	4741	TTCAGCCGGAAGTTTACACAGAAATTCATTATCTTTCTTCTAGTTACTTTATGTTTTCT	4800
Db	4741	TTCAGCCGGAAGTTTACACAGAAATTCATTATCTTTCTTCTAGTTACTTTATGTTTTCT	4800

Db 4741 TCACGCCGAGTTTACAGAAATTCATTACTTTCTTAGTTACTTATGTTTTCT 4800
QY 4801 TTTTCATTTACTCTTTCACTCACTAGTGGAAATTTATTTGTATTTTACAAATACCCCGG 4860
Db 4801 TTTTCATTTAACTCTTTCACTCACTAGTGGAAATTTATTTGTATTTTACAAATACCCCGG 4860
QY 4861 TCCATTATTTAGATTTCTTTCTCTGATGTTTGAATAGTGCATGATATATATAG 4920
Db 4861 TCCATTATTTAGATTTCTTTCTCTGATGTTTGAATAGTGCATGATATATATAG 4920
QY 4921 ATCTCAGAAATCTGAAATCTTTCTGTTCTTAATCTTTAAATATCATGTTTCTTAT 4980
Db 4921 ATCTCAGAAATCTGAAATCTTTCTGTTCTTAATCTTTAAATATCATGTTTCTTAT 4980
QY 4981 CTATCTTTTCTTATTTGT 5000
Db 4981 CTATCTTTTCTTATTTGT 5000

RESULT 2

US-10-224-562-3
; Sequence 3, Application US/10224562
; Patent No. 6730506
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C0001098DIV
; CURRENT APPLICATION NUMBER: US/10/224,562
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 53332
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-224-562-3

Query Match 100.0%; Score 5000; DB 4; Length 53332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TATAGCCCAATGCTGTGCTCAACGGGTATTTCCAGACATTTGGAGGCGAGAGATCG 60
Db 1 TATAGCCCAATGCTGTGCTCAACGGGTATTTCCAGACATTTGGAGGCGAGAGATCG 60
QY 61 CTGAGCTCAGGAATTTGAGCAAGCCCTACGTAATAGTAACTCTGTCTGTACAAA 120
Db 61 CTGAGCTCAGGAATTTGAGCAAGCCCTACGTAATAGTAACTCTGTCTGTACAAA 120
QY 121 TAAATAAAGAAATTTTCAAGCATGTGTGCGTCAACCCCAAGTGCAGCTATTTGGAGGC 180
Db 121 TAAATAAAGAAATTTTCAAGCATGTGTGCGTCAACCCCAAGTGCAGCTATTTGGAGGC 180
QY 181 TGAGGTAGAGAAATGCTTTGAAGCCAGAGATTTGAAGACAAGCTTAGGCAACATAGTGA 240
Db 181 TGAGGTAGAGAAATGCTTTGAAGCCAGAGATTTGAAGACAAGCTTAGGCAACATAGTGA 240
QY 241 CCCTGTCTATTAATAAATTAATTAAGCTGTGTCTTTGCAACAGGCTTCAGCTAGCTACT 300
Db 241 CCCTGTCTATTAATAAATTAATTAAGCTGTGTCTTTGCAACAGGCTTCAGCTAGCTACT 300
QY 301 CGAAGAACTGAGGTGGAGATCACTGAGCCAGAGAGGCTGAGGCTGCAATGAAAGTGA 360
Db 301 CGAAGAACTGAGGTGGAGATCACTGAGCCAGAGAGGCTGAGGCTGCAATGAAAGTGA 360
QY 361 TCAACCACTGATTTCCAGCTGGAAGACAGAGGAGACCTGTTTCCAAAAAATTTTAA 420
Db 361 TCAACCACTGATTTCCAGCTGGAAGACAGAGGAGACCTGTTTCCAAAAAATTTTAA 420
QY 421 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTGATG 480

Db 421 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 480
QY 481 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 480
Db 481 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 480
QY 541 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 600
Db 541 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 600
QY 601 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 660
Db 601 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 660
QY 661 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 720
Db 661 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 720
QY 721 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 780
Db 721 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 780
QY 781 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 840
Db 781 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 840
QY 841 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 900
Db 841 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 900
QY 901 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 960
Db 901 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 960
QY 961 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1020
Db 961 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1020
QY 1021 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1080
Db 1021 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1080
QY 1081 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1140
Db 1081 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1140
QY 1141 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1200
Db 1141 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1200
QY 1201 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1260
Db 1201 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1260
QY 1261 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1320
Db 1261 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1320
QY 1321 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1380
Db 1321 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1380
QY 1381 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1440
Db 1381 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1440
QY 1441 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1500
Db 1441 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1500
QY 1501 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1560
Db 1501 AAAAAAATGCAAGAAAGACATCAATAAATTTGACTGGGACATTAATTTATGTATG 1560

QY	1561	GCAGTAGATATTTTCAACTTTTTCCTGTTTAAAGATTCTGAGGTTCTTGACATCTAA	15620
Db	1561	GCAGTAGATATTTTCAACTTTTTCCTGTTTAAAGATTCTGAGGTTCTTGACATCTAA	16220
QY	1621	AAGTTAGCTAGTAGCAAAATTTGTTTTAAAGCTGAAAAACAAATGCTTTCAATAAAA	16800
Db	1621	AAGTTAGCTAGTAGCAAAATTTGTTTTAAAGCTGAAAAACAAATGCTTTCAATAAAA	16800
QY	1681	GGTAGGAGAAAAATPACTCTCAACATGTCACCTTTGACACAGSAAAACCTAATATCAA	17400
Db	1681	GGTAGGAGAAAAATPACTCTCAACATGTCACCTTTGACACAGSAAAACCTAATATCAA	17400
QY	1741	TATCACCATCAATGATATCATATATAATCATTTGATAGATAGCAATGCAATCCCTAA	18000
Db	1741	TATCACCATCAATGATATCATATATAATCATTTGATAGATAGCAATGCAATCCCTAA	18000
QY	1801	AAACTATGTATACCAATATGACCTAACTTTGSGCCAGAAACAAGACTTTAATCTTGCCAAA	18600
Db	1801	AAACTATGTATACCAATATGACCTAACTTTGSGCCAGAAACAAGACTTTAATCTTGCCAAA	18600
QY	1861	TTTTATTCTATTCAATTAACAGCTGCTGTTTTGAGTTGTGACATCTGAATGCAAGCA	19280
Db	1861	TTTTATTCTATTCAATTAACAGCTGCTGTTTTGAGTTGTGACATCTGAATGCAAGCA	19220
QY	1921	TCCCTGTCTGATGTGAGATTCTTGCACTGATAGAAAAAATGCTGAAGTTGTGAGGCT	19800
Db	1921	TCCCTGTCTGATGTGAGATTCTTGCACTGATAGAAAAAATGCTGAAGTTGTGAGGCT	19800
QY	1981	GCTCCAGGACAGACATCATATGATGCTCATAGAAAACCTCCACGCTGCACTCTGCGCA	20400
Db	1981	GCTCCAGGACAGACATCATATGATGCTCATAGAAAACCTCCACGCTGCACTCTGCGCA	20400
QY	2041	AAAGGAGAGAAACAAGATAGAGAAAGCAGTCGAGGAAAAGGTTCAAGTCAGGCCGTTTTCTC	21000
Db	2041	AAAGGAGAGAAACAAGATAGAGAAAGCAGTCGAGGAAAAGGTTCAAGTCAGGCCGTTTTCTC	21000
QY	2101	CTTGAACTTACAAAGATTATGAGGTCAGAGCTGTGTGCAAAAGCTGTACAGCTTGATGCA	21600
Db	2101	CTTGAACTTACAAAGATTATGAGGTCAGAGCTGTGTGCAAAAGCTGTACAGCTTGATGCA	21600
QY	2161	GCTGCTACCAATTGTCAGAGGAGGCGAGCAACAGCTGTGCGAGAGTGACAGCTGAGA	22200
Db	2161	GCTGCTACCAATTGTCAGAGGAGGCGAGCAACAGCTGTGCGAGAGTGACAGCTGAGA	22200
QY	2221	CCACGAGGCGCTTCAGAGGATTTGTATCCACAGAGAGACATCTTTCTCTATTGATA	22800
Db	2221	CCACGAGGCGCTTCAGAGGATTTGTATCCACAGAGAGACATCTTTCTCTATTGATA	22800
QY	2281	AAACCAAGAGTTCAACACTCCCTTTTTGTAGCGGATCTGAATCTTTGCGGTAAGTCT	23400
Db	2281	AAACCAAGAGTTCAACACTCCCTTTTTGTAGCGGATCTGAATCTTTGCGGTAAGTCT	23400
QY	2341	AAACCAATTAATAATTAATAATCTATTAAAGTCACAGAAAAATTATGCGCTGTAGTTATCAA	24000
Db	2341	AAACCAATTAATAATTAATAATCTATTAAAGTCACAGAAAAATTATGCGCTGTAGTTATCAA	24000
QY	2401	TTTGGGGAATTTCTTTGTAAACCAAAAGGAAAAAATATCTTTGGCTTTGGGCTGCACGA	24600
Db	2401	TTTGGGGAATTTCTTTGTAAACCAAAAGGAAAAAATATCTTTGGCTTTGGGCTGCACGA	24600
QY	2461	ACTCACTTGGCTTGAAGTCGAGAAAGTAGTTCTCTCAAAATCTCTAGAGTCTTAATATAC	25200
Db	2461	ACTCACTTGGCTTGAAGTCGAGAAAGTAGTTCTCTCAAAATCTCTAGAGTCTTAATATAC	25200
QY	2521	AGAGCTGAAACTTAAAGGCAAGCTGCAGATTAGTTGGTATGCTATGATTTGAACTT	25800
Db	2521	AGAGCTGAAACTTAAAGGCAAGCTGCAGATTAGTTGGTATGCTATGATTTGAACTT	25800
QY	2581	TAGTAAATTAAGTCAATGATTAATTAGCAAGCCATAGATTAATCCCTACAGCAATAAATTA	26400
Db	2581	TAGTAAATTAAGTCAATGATTAATTAGCAAGCCATAGATTAATCCCTACAGCAATAAATTA	26400

QY	2644	AGGGAATGAAAAAAGGAGGAGCTTAAACAGAAAAAGTTGCAAAACATCCATCAAA	2700
Db	2644	AGTGGACATGAAAAAAAAAGCCGACCTTAAACAGAAAAAGTTGCCAAACATCAACAA	2700
QY	2701	AGAGATTAGATTAACTCTGATGTTTAAAGACACTTTTAAAGTGAAGAAAGATGTAGTA	2760
Db	2701	AGAGATTAGATTAACTCTGATGTTTAAAGACACTTTTAAAGTGAAGAAAGATGTAGTA	2760
QY	2761	TTTTCAGGAGTTGATACATTATGCTCTTTTCAGGGATCTTTCAAGAAAAGTGCTTTTG	2820
Db	2761	TTTTCAGGAGTTGATACATTATGCTCTTTTCAGGGATCTTTCAAGAAAAGTGCTTTTG	2820
QY	2821	GGGGTACAGGAGCTTGAAGAACTTTGAAGAGGAAAAAGAGGCAATTAAGAAAAAT	2880
Db	2821	GGGGTACAGGAGCTTGAAGAACTTTGAAGAGGAAAAAGAGGCAATTAAGAAAAAT	2880
QY	2881	GGTTTACGAGGCACTGAAATCTTTACCTTGCAATTAATTTATTTCTGCTCTTTCTTTT	2940
Db	2881	GGTTTACGAGGCACTGAAATCTTTACCTTGCAATTAATTTATTTCTGCTCTTTCTTTT	2940
QY	2941	CTCTAGCTAAACAGACTTAAAGAGCATGTTTCATGACTTGACTTCAACCCGAGGGA	3000
Db	2941	CTCTAGCTAAACAGACTTAAAGAGCATGTTTCATGACTTGACTTCAACCCGAGGGA	3000
QY	3001	CTTTCAGGAGGAGTCTGAGGAGGTTTTCAGTGGGTGAGTGAGCACTGATTTGATC	3120
Db	3001	CTTTCAGGAGGAGTCTGAGGAGGTTTTCAGTGGGTGAGTGAGCACTGATTTGATC	3120
QY	3061	TTTTCAGGAGGAGTCTGAGGAGGTTTTCAGTGGGTGAGTGAGCACTGATTTGATC	3180
Db	3061	TTTTCAGGAGGAGTCTGAGGAGGTTTTCAGTGGGTGAGTGAGCACTGATTTGATC	3180
QY	3121	AAGAGAAATTTAATGTTAGCTTGTCTAACGAGGCGGACCTTGCTTCAGGGCAATTA	3240
Db	3121	AAGAGAAATTTAATGTTAGCTTGTCTAACGAGGCGGACCTTGCTTCAGGGCAATTA	3240
QY	3181	GAGGAGGCTTCCCAATGCTCTGCGCAATGCTGTAAATTTCCCTGGGGAAAAAAGT	3300
Db	3181	GAGGAGGCTTCCCAATGCTCTGCGCAATGCTGTAAATTTCCCTGGGGAAAAAAGT	3300
QY	3241	CAACACTTAAAAAAGTGTCTTTCTCTCTCTTCAACCGGCTCTTTTCCCATTTCC	3360
Db	3241	CAACACTTAAAAAAGTGTCTTTCTCTCTCTTCAACCGGCTCTTTTCCCATTTCC	3360
QY	3301	CCTAGAGCAGAGGAGAGGCTCTCCCTTTTGGGGGACCTCACTTTTACTTGAACCTTGA	3360
Db	3301	CCTAGAGCAGAGGAGAGGCTCTCCCTTTTGGGGGACCTCACTTTTACTTGAACCTTGA	3360
QY	3361	GCTGGGTGAAGGCTCTTATGCGACAGTTTAAAGGGGATTAAGCAGGTGAGTACATTA	3420
Db	3361	GCTGGGTGAAGGCTCTTATGCGACAGTTTAAAGGGGATTAAGCAGGTGAGTACATTA	3420
QY	3421	CTGGGAGAGACTTTAGAGATGAGAGTCCCGCCCCCAATTTTCAATTAATTAAGCCAG	3480
Db	3421	CTGGGAGAGACTTTAGAGATGAGAGTCCCGCCCCCAATTTTCAATTAATTAAGCCAG	3480
QY	3481	GAGACATCATAGAAAGTTCAATGACCTCAGAGCCTGTGCAAGACCACTGAGCCGAC	3540
Db	3481	GAGACATCATAGAAAGTTCAATGACCTCAGAGCCTGTGCAAGACCACTGAGCCGAC	3540
QY	3541	GAGAGACATGATTAATTAACAGCTTTGAAGAAAAAACAACCTGCGCTGCTTAATTA	3600
Db	3541	GAGAGACATGATTAATTAACAGCTTTGAAGAAAAAACAACCTGCGCTGCTTAATTA	3600
QY	3601	AATTCAGGCCACTTAAATGTTTATCAAGCTTTCCCTTCATTCATTAAGAAATTA	3660
Db	3601	AATTCAGGCCACTTAAATGTTTATCAAGCTTTCCCTTCATTCATTAAGAAATTA	3660
QY	3661	AAGAAATTAAGCAATTCCTACTACTGACCAAGAAACAATTAATCACTTCAGGCTGTG	3720
Db	3661	AAGAAATTAAGCAATTCCTACTACTGACCAAGAAACAATTAATCACTTCAGGCTGTG	3720
QY	3721	GGAGGCAACAGTGGTAAAGGCTCTTAACAGGTTTTTTTATATCCCTCCCTTAATCA	3780

Db 3721 GGAGGACAGTGGTAAAGGCTCTTAACAGGTTTTTATATCCCTCCCTAAACACAAT 3780
Qy 3781 GACAGAGTTTGTATGGCAACTGGAATTTGCTGCTTCAATTCCTCCACCGGCTTAT 3840
Db 3781 GACAGAGTTTGTATGGCAACTGGAATTTGCTGCTTCAATTCCTCCACCGGCTTAT 3840
Qy 3841 AGAAGAACTGGAAGTTGGTTTCTGCAATTATGCTACATGCAAAAGATGATAAATCTAG 3900
Db 3841 AGAAGAACTGGAAGTTGGTTTCTGCAATTATGCTACATGCAAAAGATGATAAATCTAG 3900
Qy 3901 ATTTTATATTTTGCATAATACATAAATGTCGGAATTAATAATCTGTTATCCAAA 3960
Db 3901 ATTTTATATTTTGCATAATACATAAATGTCGGAATTAATAATCTGTTATCCAAA 3960
Qy 3961 AGCTAAGTACTAATTTTGTAAACCAACTTTGTAAATATATGTAAAGATCCATGA 4020
Db 3961 AGCTAAGTACTAATTTTGTAAACCAACTTTGTAAATATATGTAAAGATCCATGA 4020
Qy 4021 ATTCCCTTTTATGCAAGTGGGAAAGTTGATGATGCTTTTCTTTATGTTATCTCA 4080
Db 4021 ATTCCCTTTTATGCAAGTGGGAAAGTTGATGATGCTTTTCTTTATGTTATCTCA 4080
Qy 4081 ATAGAGAGAAAGTAATGCTCAATAGTGTAAATATATATTTTAAATATATGAT 4140
Db 4081 ATAGAGAGAAAGTAATGCTCAATAGTGTAAATATATATTTTAAATATATGAT 4140
Qy 4141 CCGAGTGCAGTGTGTTTAACTACTTGTATCACAACCACTTACAGATTTCTTGTCT 4200
Db 4141 CCGAGTGCAGTGTGTTTAACTACTTGTATCACAACCACTTACAGATTTCTTGTCT 4200
Qy 4201 TCTCCACTCCCACTGCTTCACTTAACCTGGCCAAAAGAAAAGAAAATTTTATATA 4260
Db 4201 TCTCCACTCCCACTGCTTCACTTAACCTGGCCAAAAGAAAAGAAAATTTTATATA 4260
Qy 4261 CTACTACAGACTTAATATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 4320
Db 4261 CTACTACAGACTTAATATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 4320
Qy 4321 ACTTGTAAAGAGATTTGAGGAGATACAGTTTATTTTATTTATTTATTTATTTAT 4380
Db 4321 ACTTGTAAAGAGATTTGAGGAGATACAGTTTATTTTATTTATTTATTTATTTAT 4380
Qy 4381 TTATTTATTTATTTGATGATGATCTCTGCAACCAAGTGAAGTGAAGTGAAGTGA 4440
Db 4381 TTATTTATTTATTTGATGATGATCTCTGCAACCAAGTGAAGTGAAGTGAAGTGA 4440
Qy 4441 TCAAGTACTGCAACTCTGCTCTGAGTTCAAGCAACTCTGCTGCTGCTGCTGCTT 4500
Db 4441 TCAAGTACTGCAACTCTGCTCTGAGTTCAAGCAACTCTGCTGCTGCTGCTGCTT 4500
Qy 4501 TACTTCTTAAATCTGTTTATGATGATGATCTTAACTTTTCAATGCTGGAACATC 4560
Db 4501 TACTTCTTAAATCTGTTTATGATGATGATCTTAACTTTTCAATGCTGGAACATC 4560
Qy 4561 TGCATAAAGAGCACTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4620
Db 4561 TGCATAAAGAGCACTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4620
Qy 4621 AACTGTTTTTAAAGTGGCTGGAACGGGGGCACTTGAACATATCTTGAATCATG 4680
Db 4621 AACTGTTTTTAAAGTGGCTGGAACGGGGGCACTTGAACATATCTTGAATCATG 4680
Qy 4681 CCATGACATTCGATTCCTCTCTGAGTCTGATGATGATGATGATGATGATGATGAT 4740
Db 4681 CCATGACATTCGATTCCTCTCTGAGTCTGATGATGATGATGATGATGATGATGAT 4740
Qy 4741 TCCAGCCGGAAGTTTACAGAGATTCATTTTCTTCTTCTTCTTCTTCTTCTTCTT 4800
Db 4741 TCCAGCCGGAAGTTTACAGAGATTCATTTTCTTCTTCTTCTTCTTCTTCTTCTT 4800
Qy 4801 TTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4860
Db 4801 TTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4860

Db 4801 TTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4860
Qy 4861 TCCATTTATTAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4920
Db 4861 TCCATTTATTAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4920
Qy 4921 ATCTCAAGATTAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4980
Db 4921 ATCTCAAGATTAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4980
Qy 4981 CTATCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5000
Db 4981 CTATCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5000

RESULT 3
US-09-679-299A-18
; Sequence 18, Application US/09679299A
; Patent No. 656135
; GENERAL INFORMATION:
; APPLICANT: Vickie L. Brown-Driver
; APPLICANT: Hong Zhang
; APPLICANT: Andrew T. Wale
; TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 6 EXPRESSION
; FILE REFERENCE: RTS-0187
; CURRENT APPLICATION NUMBER: US/09/679, 299A
; CURRENT FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 164
; SEQ ID NO 18
; LENGTH: 17000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-679-299A-18

Query Match 6.7%; Score 336.2; DB 4; Length 17000;
Best Local Similarity 68.6%; Pred. No. 9.8e-63;
Matches 555; Conservative 0; Mismatches 238; Indels 16; Gaps 6;

Qy 120 ATATATTAAGAAATTTTCCAGCATGCTGCTGCAACCCCACTGCACTATTTGGAGG 179
Db 6897 ATTAATAATCTTTTGGCCAGGTGTGTGTCATGCTGTATCCACACTTTGGGAAG 6956
Qy 180 CTGAGGTGAGAGAAATGCTTGAAGCCAGAGTTGAAGCAAGCCTTGAAGCATATGAG 239
Db 6957 CCAAGGTGAGAGCACTTAAAGCAAGAGTTGAAGCAAGTCTGGGCAAGAGTGAG 7016
Qy 240 ACCCTGTCTATTAATAATATATAGTGTGTTCTTGGCAAGGCTGCAAG--CTAGCT 297
Db 7017 ACTGTGTCTTACAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7076
Qy 298 ACTGGAAAGCTGAGGTGGAGATCAC-TGAGCCAGAGGCTGAGGCTGAGTGAACA 356
Db 7077 ACTTGGAGGCTGAAGGGGAGAGATTTGTAATCCAGAGTTCAAGGCTGCAAGGAGCT 7136
Qy 357 GTGATCACCCAGCTGATTTCCAGCTGGAAGCAGAGGAGAGCCTGTTCAAAAAAA 416
Db 7137 GTGATCACCCAGCTGATTTCCAGCTGGAAGCAGAGGAGAGCCTGTTCTAAAAATAA 7196
Qy 417 AAAAAAATAATGCAAGAAAGCATCATTAACCTTGAACCTGGAACATTAATTTATGT 476
Db 7197 AAAAAAATAATTTT-----TTTCCATGTTTCCAAATGAAGATTAATAATAAT 7252
Qy 477 GATGAATTTCAACCTTTTGAAGAAATTAAGCATTTCTGATTAATAATGATATTAATA 536
Db 7253 TAAAAAATTCAGACTTCTGATGTAAGGAAAGTTACTTAT-----CATTAAGTTGTC 7306
Qy 537 TATTTATTAATTAATTAATTAATTAATTAATTTCTGAGAAACATGACTTCTGAG 596
Db 7307 AACACTATAGGTAAACAAGATCAAGATGTTGAGACATGGGCTTCTTTATGAGAG 7366
Qy 597 TTGTGATCAAAACTTTAATGATGTTTGGCGGGGTGCGGTGCTTCAAGCCTGTAAATCCA 656
Db 7367 CTTCCCTTAAGAACTTGCTG--AGTTGGCGGGGATGTTGGCTCACTCTGTATATCCG 7424

```

QY 657 GCACCTTGGAGGCGGAGGCGGTGATCAAGAGTTAGAGATCGAGACCATCTGGCT 716
DB 7425 GCACTTTGGAGGCTGAGGCGGTGATCATTAAGTCAGAGATTGACATCTGGCT 7484
QY 717 AACAGGTTAAACCTGCTCTACT--AAAAATCAAAAAATTGCGGGGTGGGTGCCAG 775
DB 7485 AACATGTTAAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7544
QY 776 AGCGCTGTAGTCCAGCTGCTCAGAGGCTGAGGCGAGAGATGTGTGAAACCGGAGG 835
DB 7545 GAGCGTGTAGTCCAGCTGCTCAGAGGCTGAGGCGAGAGATGTGTGAAACCGGAGG 7604
QY 836 CGAGCTTGCAGTGAAGCCGAGATTCGCACTGCTCAGCTGCGGCGAGTGGAGA 895
DB 7605 CGAGCTTGCAGTGAAGCCGAGATTCAGAGCCTGCTGCGGCGAGTGGAGTGGAGC 7664
QY 896 CTCTGTCTCAAAAAAATTTTCAAAAAAATTTTCAAAAAAATTTTCAAAAAAATTTT 924
DB 7665 CTCTGTCTCAAAAAAATTTTCAAAAAAATTTTCAAAAAAATTTTCAAAAAAATTTT 7693

RESULT 4
US-09-797-906-3/C
; Sequence 3, Application US/09797906
; Patent No. 6329188
; GENERAL INFORMATION:
; APPLICANT: Zianghe YAN, Karen A. KETCHUM, Valentina DIFRANCESCO, Ellen M. BEASLEY
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: US9797906
; FILE REFERENCE: CLO01151CIP
; CURRENT APPLICATION NUMBER: US/09/797, 906
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 84495
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(84495)
; OTHER INFORMATION: n = A,T,C or G
US-09-797-906-3

Query Match 6.7%; Score 334.4; DB 3; Length 84495;
Best Local Similarity 61.8%; Pred. No. 4.4e-62;
Matches 656; Conservative 0; Mismatches 386; Indels 20; Gaps 7;

QY 119 AATAATTAAGAATTTTCAGAGCATGTGCGTGCACCCCACTGCGAGTATTTGGAG 178
DB 38013 AATAATTAAGAATTTTCAGAGCATGTGCGTGCACCCCACTGCGAGTATTTGGAG 37954
QY 179 GCTGAGTGAAGAGTCTTGAAGCAGAGTGAAGAGACCCCTGAGCAATGTGA 238
DB 37953 ACAGAGTGAAGAGTCTTGAAGCAGAGTGAAGAGACCCCTGAGCAATGTGA 37894
QY 239 GACCTGTGTCTATAAAT--AAATTAGTGTGTTGTTGGCA--CAGGCTGAGC 292
DB 37893 GACCTGTGTCTATAAATTTTAAATTTAGCCAGGAGTGTGCTGCTGTGTGTC 37834
QY 293 TAGCTACTCGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 351
DB 37833 CAGCTACTCGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 37774
QY 352 GAACTGTATCAACCACTGAGTTCAGAGCTGGAAGAGAGAGAGAGAGAGAGAGAG 411
DB 37773 GAGCTGTATCAACCACTGAGTTCAGAGCTGGAAGAGAGAGAGAGAGAGAGAGAG 37715
QY 412 AAAAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 471
DB 37714 CAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 37658

```

```

QY 472 TATGATGAAATTCAGATCTTTTAGAGAGAAATTTAGCATTTCTGATAAATGTATAT 531
DB 37657 AAGATTAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 37599
QY 532 AATTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 591
DB 37598 ----AAGATTTTGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 37543
QY 592 CTGATTTTCTTAAAAACAGATCAG--TAGGCGGGGGGGGGGGGGGGGGGGGGGGGG 651
DB 37542 CTGATTTTCTTAAAAACAGATCAG--TAGGCGGGGGGGGGGGGGGGGGGGGGGGGG 37487
QY 652 TCCAGCACTTTGGAGGCGGAGGCGGTGATTCAGAGAGAGAGAGAGAGAGAGAG 711
DB 37486 TCCAGCACTTTGGAGGCGGAGGCGGTGATTCAGAGAGAGAGAGAGAGAGAGAG 37427
QY 712 TGGCTTAACAGGAGTGAACCTGCTCTAATAAATTAAGCAAAAAATTAAGCGGGGTGG 771
DB 37426 TGGCTTAACAGGAGTGAACCTGCTCTAATAAATTAAGCAAAAAATTAAGCGGGGTGG 37367
QY 772 CCAGAGCGCTGTAGTCCAGCTGCTCAGAGGCTGAGGAGAGAGAGAGAGAGAGAG 831
DB 37366 GCGGCGGCTGAGTCCAGCTGCTCAGAGGCTGAGGAGAGAGAGAGAGAGAGAGAG 37307
QY 832 GAGGCGGAGCTTGCAGTGAAGCCGAGATTCGCACTGCTCAGCTGCGGCGAGCATGC 891
DB 37306 GAGGCGGAGCTTGCAGTGAAGCCGAGATTCGCACTGCTCAGCTGCGGCGAGCATGC 37247
QY 892 GAGACTGTCTCAAAAAAATTTTCAAAAAAATTTTCAAAAAAATTTTCAAAAAAATTTT 951
DB 37246 GAGACTGTCTCAAAAAAATTTTCAAAAAAATTTTCAAAAAAATTTTCAAAAAAATTTT 37187
QY 952 TTCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1011
DB 37186 GTTACATGATTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 37127
QY 1012 GAATGAGCTTAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1071
DB 37126 TCTCTACATATTTCAAAATTAATTTGAGAGTGAAGAGAGAGAGAGAGAGAG 37067
QY 1072 TTAATTTACGTAACATTTTAAAGATCATGAGTGAAGTGAAGTGAAGTGAAGT 1131
DB 37066 AGAATTTAAACACAAAAAATTTTAAAGATCATGAGTGAAGTGAAGTGAAGTGA 37007
QY 1132 TATGTTGTGATTAATGACATTAATTTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1173
DB 37006 TATTTGAAGATTAATTTTAAATCTTTATTTGATGTGTTA 36965

RESULT 5
US-09-146-053-6
; Sequence 6, Application US/09146053A
; Patent No. 639349
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; APPLICANT: Sprinkle, Terry Joe Curtis
; APPLICANT: Venema, Richard C.
; TITLE OF INVENTION: Human Amino-peptidase P Gene
; FILE REFERENCE: MCG103
; CURRENT APPLICATION NUMBER: US/09/146, 053A
; CURRENT FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/057,854
; EARLIER FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 45546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-053-6

Query Match 6.7%; Score 333.8; DB 3; Length 45546;

```

Best Local Similarity 66.7%; Pred. No. 4,7e-62;
Matches 540; Conservative 0; Mismatches 257; Indels 12; Gaps 4;
QY 136 CCAGGACATGGTGGCGTGCACCCCGAGTCCAGCTATTGGAGGCTGAGGAGGAT 195
Db 36453 CCGGGTGCATGGCTTACCTGTAATTGACGACCTTTGGAGGTAGAGGAGATC 36512
QY 196 GCTTGAAGCAGAGATTGAAGACAGCTTAGGCAACATAGTGAAGCCCTGTCTATMAA 255
Db 36513 ACTTGAGGCCAGAGATTCAAGATCAGCTGAGCAACATAGGCAAA-----AAAAAAA 36566
QY 256 AATTAATTAAGTGGTGTCTTTGGACAGGCTGAG--CTAGCTACTCGGAATCGAGG 313
Db 36567 AAAAATTTAGCCAGAAATGGTGGCATGCTCTGTAATCCAACTACCCGGAGGCTGAGG 36626
QY 314 TGGAGAGATCATGAGCCCGAGAGGCTGAGGCTGAGTGAACAGTATCAACCAAGCTGA 373
Db 36627 TGGAGAGATCATGAGCCCGAGAGGTTGAAGCTGAGCGAGCTGTGACACCACTGCA 36686
QY 374 TTCCAGGCTTGAAGACAGAGGAGACCTGTCTTCCAAAAAATTTAAATTTAA 433
Db 36687 TTCCAGGCTTGAAGACAGAGGAGACCTGTCTTCAATTTAATTTAATTTAATTTA 36746
QY 434 AGAAGACATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 493
Db 36747 ATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 36803
QY 494 TTTAGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 553
Db 36804 CTATCCAACTCTTATCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 36863
QY 554 ATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 613
Db 36864 TTGTGACATGCTGATCTTACCTGCCCCAAATATCCACTATGATCATCTAATAAT 36923
QY 614 AATGCTCT--TTGGCCGGGTCGGGTGAGCTCAGCCTGTATATCCAGACCTTTGGAGGCG 672
Db 36924 AAGAATTTCTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 36983
QY 673 AGCGGGGTGATCAACAAGTTAGAGATGAGACCATCTGCTTAACAGCGTAAACCTTC 732
Db 36984 AGGTGGGAGATCAAGAGATCAAGAGATCAAGACCATCTGCTTAACATGATGATG 37043
QY 733 GTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 792
Db 37044 GTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 37103
QY 793 TGCTCAGAGGCTGAGGAGAGATGATGATGATGATGATGATGATGATGATGATG 852
Db 37104 TACCGAGAGGCTGAGGAGAGATGATGATGATGATGATGATGATGATGATGATG 37163
QY 853 CGAGATTCGCTCACTGATCCAGCTGAGGAGCAAGTGCAGACTGTGTCTCAAAAAA 912
Db 37164 CGAGATTCGCTCACTGATCCAGCTGAGGAGCAAGATGATGATGATGATGATGAT 37223
QY 913 AAAAAAAAAAATGATGCTTTGAG 941
Db 37224 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 37252

RESULT 6
US-09-801-052-3
; Sequence 3, Application US/09801052
; Patent No. 6368842
; GENERAL INFORMATION:
; APPLICANT: BRASLEY, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001045
; CURRENT APPLICATION NUMBER: US/09/801, 052
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 5

SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 16063
; TYPE: DNA
; ORGANISM: Human
US-09-801-052-3
Query Match
Best Local Similarity 67.7%; Pred. No. 4.5e-59;
Matches 540; Conservative 0; Mismatches 243; Indels 15; Gaps 6;
QY 136 CCAGGACATGGTGGCGTGCACCCCGAGTCCAGCTATTGGAGGCTGAGGAGGAT 195
Db 10451 CCGGGGTGATGCTTACCTGATCCAACTTTTGGAGGTGAGGAGGATC 10510
QY 196 GCTTGAAGCAGAGATTGAAGACAGCTTAGGCAACATAGTGAAGCCCTGTCTATAA 255
Db 10511 ACTTGAGCTCAGAGATTGAAGACAGCTTAGGCAACATAGTGAAGCCCTGTCTATA 10570
QY 256 AAT-----AATTAAGTGGTGTCTTTGGACAGGCTGCA--CTAGCTACTCGAAGAC 308
Db 10571 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10630
QY 309 TGAGGTGGAGGATGATC--TGAGCCAGAGGCTGAGGCTGAGTGAACAGTATCAACCA 367
Db 10631 TGAGGTGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 10690
QY 368 GCTGATTCACCTGAGAGACAGAGGAG--ACCTGTTTCCAAAAAATTTAAATTTAA 425
Db 10691 ACTGCTCTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 10750
QY 426 AATATGCAAGAAACATCATTAATTAATTAATTAATTAATTAATTAATTAATTA 485
Db 10751 AATATGCAAGAAACATCATTAATTAATTAATTAATTAATTAATTAATTAATTA 10810
QY 486 CAAATCTTTTGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 545
Db 10811 TGTGGATTTGGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGA 10870
QY 546 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 605
Db 10871 GGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10928
QY 606 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 665
Db 10929 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10985
QY 666 GAGGCGAGGCGGATGATCAAGAGTTAGAGATGAGACCATCTGCTTAACAGGATG 725
Db 10986 GAGGCGAGGCGGATGATCAAGAGTTAGAGATGAGACCATCTGCTTAACAGGATG 11045
QY 726 AATCTGCTCTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 785
Db 11046 AATCTGCTCTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 11105
QY 786 TCCAGCTCTCAGAGGCTGAGGAGAGATGATGATGATGATGATGATGATGATGATG 845
Db 11106 TCCAGCTCTCAGAGGCTGAGGAGAGATGATGATGATGATGATGATGATGATGATG 11165
QY 846 AGTGAAGCAGAGTTGCGCACTGCACTCAAGCTGAGGAGAGAGTGAAGCTGTCTCA 905
Db 11166 AGTGAAGCAGAGTTGCGCACTGCACTCAAGCTGAGGAGAGAGTGAAGCTGTCTCA 11225
QY 906 AAAAAAAAAAAAAAAAAA 923
Db 11226 AAAAAAAAAAAAAAAAAA 11243

RESULT 7
US-10-020-121-3
; Sequence 3, Application US/10020121
; Patent No. 6638747
; GENERAL INFORMATION:

APPLICANT: BEASLEY, Ellen et al.
 TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEIN,
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE
 TITLE OF INVENTION: PROTEINS, AND USES THEREOF
 FILE REFERENCE: CL001045DIY
 CURRENT APPLICATION NUMBER: US/10/020,121
 CURRENT FILING DATE: 2001-12-18
 PRIOR APPLICATION NUMBER: US 60/255,386
 PRIOR FILING DATE: 2000-12-15
 PRIOR APPLICATION NUMBER: US 09/801,052
 PRIOR FILING DATE: 2001-03-08
 NUMBER OF SEQ ID NOS: 5
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 16063
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-020-121-3

Query Match 6.4%; Score 319.2; DB 4; Length 16063;
 Best Local Similarity 67.7%; Pred. No. 4,5e-59;
 Matches 540; Conservative 0; Mismatches 243; Indels 15; Gaps 6;

136 CCAGGAGTGGCGCCGCAACCCAGCTAGCTATTTGGAGGCTGAGGTAGAGGAAT 195
 10451 CCGGGCGTGGTGTCTACCGCTGTATCCACATTTGGGGGCTGAGGTGGATATC 10510
 196 GCTTGAGCCAGAGTGTGAAGCAAGCCTAGGCAATAGTAGAGCCCTGTCTATAA 255
 10511 ACTTGAGTCTAGAGTTCAGACACACCTGGCTTAACATGTGTAAATCCATCTCTTA 10570
 256 AAT-----ATTAGCTGTTGCTTGGCAAGGCTGCA--CTAGTACTCTGGAAGAC 308
 10571 AATATCAAAATTTAGCCGGCGCTGTGGGACCTGTGTATCCAGCTACTTGAAGGC 10630
 309 TAGGAGGAGAGATCAC--TGAGCCAGAGGCTGAGGCTGAGTGAACATGATCCCA 367
 10631 TAGAGGAGAGATGCTTGAACCTGGAGGTGAAGTGTGAGTGTGAGATCCGCC 10690
 368 GCTGATTCAGCCTGGAAGACAGAGGAG--ACCTGTTTCCAAAAAATTTTTTTT 425
 10691 ACTGCTCTCAGCTGTGGAGACAGACAGCAAGATCTGTCTCAAAAAAACAACAAA 10750
 426 AATATGCAAGAAAGACATCATTAATCTTGAAGCTGGAATTAATTTATGTGAATT 485
 10751 AAAAAAGAAAGCTCAGAAATTTGGGCTTGAAGGCAAGAACCTGAGGAGTGCAGA 10810
 486 CACAATCTTTAGAGAAATTAAGCATTTCTGATAAATATATTAATTAATTAAT 545
 10811 TGTGGATTGGAGAGTGAAGGAGGAGGCTGGAACAAGATGAGGGCTTGGGGATG 10870
 546 AATTTCAATGAAATTAATTAATTTCTGAGAAATAGCTTCTCACTCTCAAGTTGTGATC 605
 10871 GGGACATTAAGATTTTG--GTTTGGGATGCAAGCTTGGGATTTGTGATCTCTGGATA 10928
 606 AAAATTTAATGCTTTGGCGGGGCTGCTGCAAGCTGTAAATCCAGACATTTGG 665
 10929 AGAAGATTAACAAC--GGCGGAGCTGTGGCTCAACCTGTAAATCCAGACATTTGG 10985
 666 GAGGCGGAGCGGGTGCATCAAGAGTTAGAGATGAGACCATCTCTGCTTAAACAAGTGTG 725
 10986 GAGGCTGAGGCGGGTGTATCAGAGGTGAGAGATGAGACCATCTCTGCAACAAGTGTG 11045
 726 AATCTGCTCTCTACTAATAAATTAACAATAATTAAGCGGGTGTGCGGTCCAGACGCTGTAG 785
 11046 AAGCCCCGTCTCTACTAATAAATTAACAATAATTAAGCGGTGTGCGCGCTGTAG 11105
 786 TCCAGCTGCTCTCAAGAGCTGAGGAGAGATGTTGTAAATCCGAGGAGCGGAGCTTGC 845
 11106 TCCAGCTGCTCTGAGGAGCTGAGGAGAGATGTTGTAAATCCGAGGAGCGGAGCTTGC 11165
 846 AATGAGCGAGATGTTGGGCACTGCACTCAAGCTGGGCGACAGTGGAGATCTGTCTCA 905
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 11166 AGTAGCCCAAGATCATGCCACTGCACTCGAGCTGGGCGACGAGAGGAGACACCTTCTCA 11225
 Oy 906 AAAAAAAAAAAAAAAAAA 923
 Db 11226 AAAAAAAAAAAAAAAAAA 11243

RESULT 8
 US-09-792-616-1/C

; Sequence 1, Application US/09792616
 ; Patent No. 6780587
 ; GENERAL INFORMATION:
 ; APPLICANT: PKE International, Inc.
 ; APPLICANT: University of Hawaii
 ; TITLE OF INVENTION: Mutations in a gene encoding an ABC transporter (MRP6) causing
 ; FILE REFERENCE: PKE-001
 ; CURRENT APPLICATION NUMBER: US/09/792,616
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 107820
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc. feature
 ; OTHER INFORMATION: "n" can be an A or a T or a G or a C
 US-09-792-616-1

Query Match 6.3%; Score 317; DB 4; Length 107820;
 Best Local Similarity 64.0%; Pred. No. 2.8e-58;
 Matches 620; Conservative 0; Mismatches 315; Indels 34; Gaps 8;

13 CTGTGCTCAGCGGTGTATTTCCAGACATTTGGAGG-----CAGAGAGATGCTTGAAG 66
 89369 CAGTGGCTACACCATTAATCCCAACATTTGGAGGCCAACCCCTGACAGATCATTGAG 89310
 67 CTCAGGAATTTGAGACAAAGCTTACGTAATAGTGAACCTGTGTGTACAAATTAATA 126
 89309 GCCGGAGTTGAGACACACCTGCGCCACATGAGCGAAATCCCA--TCTACTATAAACA 89251
 127 AAGATTTTTCAGGAGATGTTGGCGCTGCACCCCAAGTCCAGATTAATTTGGAGGCTGAGGT 186
 89250 AAAAAATTGACAGGTGTGTGCTTAACGCTGTAGCTTGAACATTTGGAGGCTGAGGT 89191
 187 AGGAGGAATGCTTGAAGCAGAGGTTGAAGACAGGCTTGAAGACATAGTGAACCTGT 246
 89190 GGGAGGATTGCTTGAAGCCAGTAGAGTCAAGACCAAGCTTGGCCAGATGATATACCAT 89131
 247 GTCTATAAAAAATTAATTAAGCTGTGTCTTGGACAGGCTTCAAGTACTGCTGGAAG 306
 89130 CCAAAAAACAATAAATTAAGCTGTGTATG--TGGCTGTAAATTTCAACTCAAGAG 89072
 307 ACTGAGTGGAGATCAC--TGAGCCAGAGGCTGAGGCTGCAATGAGCAAGTATCACC 365
 89071 GCTGAGGCTGTGAATTCGCTTGAAGCCAGGAGGAGGTTTCAAGTACAGATCACA 89012
 366 CAGCTGATTCAGGCTGGAAGACAGAGGAGACCCCTGTTCAAAAAAATTTTTTTT 425
 89011 TCAAGTCACTCAGAGCTTGAAGCAAGATGAGATCTCATCTCAAAAAAATTAATTAAT 88952
 426 AATATGCAAGAAAGACATCATTAATTTGACCTGGAGACATTAATTTA--TGTATGAAA 483
 88951 TAAAAAAAAGAGCTGAGGAGAGAGTCAAGCTTCTTACATATGATTTTGTGCAAC 88892
 484 TTCAATTTTATAGAGAAATTAAGATTTGTGATTAATTAATTAATTAATTTAATTT 540
 88891 GGAATTTGATTTCCAGGAGAAATGCTGTTCCTTGCACAAAGGCTATCCCTGTGTTAG 88832
 541 -----ATTATAATTCMAATGAGATTAATTAATTTCTGAGAAATGAGTT 583
 88831 CCACTGAATCTCTGCACTTCTATTTCTTCCCATCTGTAAATAGATTAATTAAT 88772
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY	584	CTCACTCTTCAGTTCGTCAGTCAAAATTAAATGCTCTTTGGCCGGTGCAGTGGCTCAC	643
Db	88771	ATACTCATACATTATATACAAATTATAACTATCAATTACAGGCTGGGCTGGTGGCTCAT	88712
QY	644	GCCCTGTAATCCCAAGCATTTTGGAGGCGCGAGCGGGTGATCAC--AAGTTAGAGATC	701
Db	88711	GCCTGTAATCCAGCACTTTGGGAGGCTGAGGAGAGGTGATCATCTGABGTCAAGAGTTTC	88652
QY	702	GAGACCATCTCTGGCTAAACAAGGTGAAAACCTGTCTCTACTAAATAATACAAAAATTAGCC	761
Db	88651	GAGACCAAGTCTGGCGCAACATGTGAAACTCATCTCTACAAAAATAC--AAAAATTAGCC	88593
QY	762	GGGAGGGGTCCAGAGGCTGTGATGCCAGCTCTCAAGAGGTGAGGCAAGAAATGAT	821
Db	88592	GGGATATGTGTGGTGTCTCTGTATCCCACTATCCAGAGGCTGAGGTGGAGAAATCC	88533
QY	822	GTGAATCCCGGAGGCGAGCTTCACTGAGAGCCAGATTTGGGCGACATCGCATCCAGGCTTG	881
Db	88532	TTGAATCTGGGAGGTGAGGTTCAGATGAGGCCAGATTTGACCACTGCACTCCAGGCTTG	88473
QY	882	GCGACAGTGCAGACTCTGTGCTCAAAAAAAAAAAAAAAAAAAGTTGAATGCTCTTTGAG	941
Db	88472	GAGACAGAAACAAGCTCATCTCAAAACAACATATATTAACCTTAATGGGTCAATGA	88413
QY	942	CCAAAGTAGT	950
Db	88412	AGAAATTAAT	88404

```

RESULT 9
US-10-118-328-3
; Sequence 3, Application US/10118328
; Patent No. 6773904
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001220
; CURRENT APPLICATION NUMBER: US/10/118,328
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/282,460
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 30350
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-118-328-3

```

Query Match	6.2%	Score 309,	DB 4,	Length 30250;
Best Local Similarity	65.4%	Pred. No.	9e-57;	
Matches 639;	Conservative	0;	Mismatches 270;	Indels 68;
				Gaps 10;

QY	15	GTGGCTCAACCGTGATATTCACAGCATTTGG-----AGGACAGAGAGATCGCTTAGGCT	68
Db	25156	GTGGCTCAACACTGTATATCCACACACTTTGGAGAGTCAAGGTGGAGATTGTCTAGCT	25215
QY	69	CAGGAATTGGAGACAAGACCTTACGTATCATATGTAACCTCTGTCTGTATCAAAATTAATATAA	128
Db	25216	CAGAGATTTTAGACCCAGCTCGGCAACATAGAGAGAC--CCATCTATACAAAAATATAGAG	25274
QY	129	GAATTTTCCAGGACAGGTGGC-GTGACACCCCGATGCGACGTATTTTGGAGGCTGAGGTA	187
Db	25275	AAATTTGGCTGGAGGCGGTGGCTCAAGCCTGTATATCCGACATTTTGGAGGCGCAAAAGG	25334
QY	188	GAGAGATGCTTTGAAGCCAGAGATTGAAGACAAGACCTTAGGCAACATATGAGACCTGTG	247
Db	25335	GGTGGATCACTTGAAGTCAAGGATTTGAGAGACAGCCTGGCCAAACATATGATGAAGCCCCGTC	25394
QY	248	TCCT-----ATAAAAATATTAAGTGTGTTGCTTTGGACACAGGCTGCAG--CTAGCTA	298

Db		25395	TCTATCAAAAATACAAAAAATTAGCTGGGCATGTTGGCATGTCTGAATGTCACGCTA	2545
OY		299	CTCGAAAGACTGAGGTGGAGGATCAC-TGAGCCAGAAGGCTGAGCCTGCAGTGAACAG	357
Db		25455	CTCGGAGAGCTGAGGCGAAGAGATTCGCTTGAAACCAGAGAGCGAGAGTTGCAATGAGCCG	25514
OY		358	TGATACCCAGCTGAGATTCAGGCTTGAAAGACAGAGAGGAGACCCCTGTTCCAAAAAAAA	417
Db		25515	AGATTCGACCATGTCATCCAGCTCTGG-CGACTGAGACAAGACTCTGTCTCAAAAAA	25573
OY		418	AAAAAAAAAAAAATGCAGAAAAAGACATCATTAACCTTGACCT-----	458
Db		25574	AAAAAAAAAAAAAAAAAGCATATATATATATATATATATATATAGAGAGAGAGAGAGA	25633
OY		459	-----GGGACATACTTTATATGTGATGAATTCACAATCTTTAGAGAAATTA	508
Db		25634	GAGAGAGAGAGAGAGAGATTTAGCTGACATGTTGGTCATGTGCTCTGATTCCCAATCC	25693
OY		509	GCATTTCTGATTAATAGATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	568
Db		25694	AACACTGCGGAGAGCTGAGTGGAGAGATCACTTGAGCCTAGAGAGTGGAGGCTGCAGCG	25753
OY		569	CTGAGAACTAGCTTCTCACTCTCTCAGTTTCAGTCAA-----A	609
Db		25754	AGCTGAGATCACGCCCATCTGACTCCAGCCTGGGGTGCACAGCAGAACCCCTGTCAATTA	25813
OY		610	CTTTAATNGTCTTTGGCCGGGTGGCTCAGCCCTGTATATCCAGACATTTGGGAG	669
Db		25814	AAAAAAAAAAAAAGGGGCCGGGCGCATGTGGCTCACGCTGTAAATCCAGACATTTGGGAGG	25873
OY		670	CCGAGCGGGTGTGATCACAAGTTTAGAATGCAGACCAATCTCTGCTAACACGCTGAAC	729
Db		25874	CCGAGGGGGTGTGATCACGAAGTCAAGAAATGGAAGCAATCTGTGGCAACATGTGAAGC	25933
OY		730	CTGTCTCTACTAAAAATACAAAAAATTAGCCGGGTGGCTGCCAGACGCTGTATGCTCC	789
Db		25934	CTGTCTCTACTAAAAATACAAAAAATTAGCCAGGATGTGTGGCAGCGCCCTATATGTCCC	25993
OY		790	AGCTGCTCAGAGGCTGAGCAGAGAAATGTGTGAACCCGGGAGGCGAGCTTGACAGTG	849
Db		25994	AGCTACTCAGAGAGGCTGAGCGAAGAAATCACTTGAAACCAGAGAGGTGAAGTTCAAGTG	26053
OY		850	AGCCGAGATTTGCCCATCTGCACTCCAGCCTGGGCGACAGTGCAGAACTTGTCTCAAAA	909
Db		26054	AGCCAGATTTGCCCATCTGCACTCCAGCCT-GGCCGACAGAGTGAATCTCGTCTCAAAA	26112
OY		910	AAAAAAAAAAAAAAGT	926
Db		26113	AAAAAAAAAAAAATGATT	26129

RESULT 10
US-09-817-180-3
; Sequence 3, Application US/09817180

```

? APPLICANT: GAN, Weiniu et al.
? TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEOTIC
? TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
? TITLE OF INVENTION: THEREOF
? FILE REFERENCE: CLO01183
? CURRENT APPLICATION NUMBER: US/09/817,180
? CURRENT FILING DATE: 2001-03-27
? NUMBER OF SEQ ID NOS: 4
? SOFTWARE: PasteSeq for Windows Version 4.0
? SEQ ID NO 3
? LENGTH: 15297
? TYPE: DNA
? ORGANISM: Human
? JS-09-817-180-3

```

Query Match	6.1%;	Score 303.6;	DB 3;	Length 15297;
-------------	-------	--------------	-------	---------------

Best Local Similarity 65.0%; Pred. No. 1e-55;
Matches 603; Conservative 0; Mismatches 289; Indels 36; Gaps 9;

```
QY 13 CTGTGGCTCAGCGGCTGATTTCCAGACACTTTGGGAGGC-----AGAGAGATCGCTTGAG 66
Db 11700 CAGTTGCTACCGCTGTGATCCAGACACTTTGGGAGGCCTGAGTGGGTGATCACTTGAG 11759
QY 67 CTCAGAAATTGGACACAGCCTACGTAACTAGTAACTCTGTCTGTAATAATA 126
Db 11760 CCCAGAGATTCAAGATCAGCTTGAGACACAGCTGAAA-CTCATCTGTAAATAATAATAC 11818
QY 127 AAGAATTTTCAGGAGCTGAGGCTGACACCCAGTGCAGCACTTATTTGGAGGCTAGGT 186
Db 11819 AAAAAATGACTGGGACAGGCTGCTCACTGTAACTCCAGACACTTTGGAGGCGGAGGC 11878
QY 187 AGAGAGATGCTTGAAGCCAGAGATTGAAGACAAAGCTAGCAATAGTGAACCTGT 246
Db 11879 AGGTGATCACCCTGTGTGATGAGATTGAGACAGCCAGCAACAACTGTGAACCCCAT 11938
QY 247 GTCTATATAAAAT-----AATTAGCTGTGTCTTGGCAGAGCTGCA--GCTAGCTAC 299
Db 11939 CTCTATATAAAATATAAAATTAAGCAGGCTGTGTGCTGTGTATCCAGACTAC 11998
QY 300 TCGGAAGATGAGAGTGGGAGGATCAC-TGAGCCAGAGGCTGAGGCTGAGTGAACAGT 358
Db 11999 TTGGAGGCTGAGAGTGGAGAAATGCTTGAACCCAGAGAGGCGAGGCTGAGTGAACCA 12058
QY 359 GATCACCCAGCTGATTCAGCCTGGAAGAC-AGAGGAGAGCCCTGTTCCTCAAAAAA 417
Db 12059 GATTGTGCCACTGCACTCCAGCCTGGGCGCAGAGATGAACTCACTCAAAAAAACCC 12118
QY 418 AAAAAAAAAAATGCAAGAAAGACATCATAACTTGACCTGGAGACATACTTTATG 477
Db 12119 AAAAAAATAAATAATCAAAATATAGCTGGGTGTGTGACATGTGCTGTATGCTGTCTA 12178
QY 478 ATGAATATCAATCTTTTGAAGAAATTAGCACTTGTATTAATTAATTAAT 537
Db 12179 CTGGGAGGCTGAGTGGAGGATCACTGAGCCGAGAGTGGAGGTGCACTGAGCTG 12238
QY 538 ATTATTAATAATCAATGGAATTAATTAATTTCTGAAGAACTAGCTTCTCACTCTCACT 597
Db 12239 A-----GATCATGCTGACATGCCACCCCACTGGGTGACAGAGAGAGA 12281
QY 598 TGTCACTCAAACTTTAATGCTTTGGCCGGGTGCGTGTGCTCAGCTGTATCCAG 657
Db 12282 GAGACCTTCACTCGAAAAAGAAAAAACTGGGCGGAGTGGCTCAGCCTGTATTTCAA 12341
QY 658 CACTTTGGGAGGCGGAGGCGGTGATCAC--AAGTTAGAGATGAGAACCATCTGCG 715
Db 12342 CATTTTGGGAGGCTGAGGAGGTGATCACTTGAATCTAGAGATTGACACTAGCTGCG 12401
QY 716 TAACACGCTGAACCTGCTCTACTATAAAATACAAAAAATTAGCCGGGTGCGGTCCAG 775
Db 12402 CAACATGCGAAGAACT-GTCTCTACTATAAAATTAAGCAAGGTGATGCTGTC 12460
QY 776 AGGCTGTAGTCCAGCTGCTCAGAGGCTGAGGCTGAGGAGATGTTGTAACCCGGAG 835
Db 12461 AAGCCTGTATCTCAACTACTTGGGAGGCTGAGGAGCAAGAAATCGCTTGAACCTGGAG 12520
QY 836 CGGAGCTTGCAGTGAAGCGAGATTGGGCACTGCACTCAGAGCTGGGCGACAGTGGAGA 895
Db 12521 TGAAGGTGCAAGTGAAGTGAATCACAACCTGCACTTCCAGCGTGGGTGACAGAGCAAG 12580
QY 896 CTCTGTCTCAAAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 923
Db 12581 CTCATCTCAAGAAAAAAGAAAAAATAATAATAATAATAATAATAATAATAATAATAATA 12608
```

RESULT 11
US-10-003-295-3
; Sequence 3, Application US/10003295
; Patent No. 6686187
; GENERAL INFORMATION:

APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THERMO
; FILE REFERENCE: CL001183DIV
; CURRENT APPLICATION NUMBER: US/10/003,295
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-003-295-3

Query Match 6.1%; Score 303.6; DB 4; Length 15297;
Best Local Similarity 65.0%; Pred. No. 1e-55;
Matches 603; Conservative 0; Mismatches 289; Indels 36; Gaps 9;

```
QY 13 CTGTGGCTCAGCGGCTGATTTCCAGACACTTTGGGAGGC-----AGAGAGATCGCTTGAG 66
Db 11700 CAGTTGCTACCGCTGTGATCCAGACACTTTGGGAGGCTGAGTGGGTGATCACTTGAG 11759
QY 67 CTCAGAAATTGGACACAGCCTACGTAACTAGTAACTCTGTCTGTAATAATA 126
Db 11760 CCCAGAGATTCAAGATCAGCTTGAGACACAGTGA-CTCATCTGTAAATAATAATAC 11818
QY 127 AAGAATTTTCAGGAGCTGAGGCTGACACCCAGTGCAGCACTTATTTGGAGGCTGAGGT 186
Db 11819 AAAAAATGACTGGGACAGGCTGTGATCACTGTAACTCCAGACACTTTGGAGGCGGAGGC 11878
QY 187 AGAGAGATGCTTGAAGCCAGAGATTGAAGACAAAGCTTAGCAATAGTGAACCTGT 246
Db 11879 AGGTGATCACCCTGTGTGATGAGATTGAGACAGCAGCAACACATGTGTGAACCCCAT 11938
QY 247 GTCTATATAAAAT-----AATTAGCTGTGTCTTGGCAGAGCTGCA--GCTAGCTAC 299
Db 11939 CTCTATATAAAATATAAAATTAAGCAGGCTGTGTGCTGTGTATCCAGACTAC 11998
QY 300 TCGGAAGATGAGAGTGGGAGGATCAC-TGAGCCAGAGGCTGAGGCTGAGTGAACAGT 358
Db 11999 TTGGAGGCTGAGAGTGGAGAAATGCTTGAACCCAGAGAGGCGAGGCTGAGTGAACCA 12058
QY 359 GATCACCCAGCTGATTCAGCCTGGAAGAC-AGAGGAGAGCCCTGTTTCAAAAAA 417
Db 12059 GATTGTGCCACTGCACTCCAGCCTGGGCGCAGAGATGAACTCAATCTCAAAAAAACCC 12118
QY 418 AAAAAAAAAAATGCAAGAAAGACATCAATACTTGACCTGGGACATACTTTATG 477
Db 12119 AAAAAAATAAATAATCAAAATATAGCTGGGTGTGTGACATGCGCTGTATGCTGTCTA 12178
QY 478 ATGAATATCAATCTTTTGAAGAAATTAGCACTTGTATTAATTAATTAAT 537
Db 12179 CTGGGAGGCTGAGTGGAGGATCACTGAGACCCGAGAGTGGAGGTGCACTGAGCTG 12238
QY 538 ATTATTAATAATCAATGGAATTAATTAATTTCTGAAGAACTAGCTTCTCACTCTCACT 597
Db 12239 A-----GATCATGCTGACATGCCACCCCACTGGGTGACAGAGAGAGA 12281
QY 598 TGTCACTCAAACTTTAATGCTTTGGCCGGGTGCGTGTGCTCAGCTGTATCCAG 657
Db 12282 GAGACCTTGACTGAAAAAAGAAAAAACTGGGCGGAGTGGCTCAGCCTGTATTTCAA 12341
QY 658 CACTTTGGGAGGCGGAGGCGGTGATCAC--AAGTTAGAGATGAGAACCATCTGCG 715
Db 12342 CATTTTGGGAGGCTGAGAGGTTGATCACTTGAATCTAGAGATTGACACTAGCTGCG 12401
QY 716 TAACACGCTGAACCTGCTCTACTATAAAATACAAAAAATTAGCCGGGTGCGGTCCAG 775
Db 12402 CAACATGCGAAGAACT-GTCTCTACTATAAAATTAAGCAAGGTGATGCTGTC 12460
QY 776 AGGCTGTAGTCCAGGCTGCTCAGAGGCTGAGGAGGAGATGTTGTAACCCGGAGG 835
```


Db 13896 TCGTTCCTGTAATAGTGGAGTGAATCACTTATCATCCAGGCTATTTTCATGCAAT 13837
Qy 561 TAAATATTCGAGAACTAGCTTCTCACTCTCTCACTTGTCACTCAAACTTTATGATC 620
Db 13836 AAAATGGAATACGCTTCTTCTTCCCACTCCAGGATTTGGAGAGAAATGTCTAACCA 13777
Qy 621 TTTGGCCGGGTGGGGTGGCTCAGCCTGTATCTCCAGCACTTTGGAGGCGGAGCGGGT 680
Db 13776 ACATATCTTAAAAACATCTGTAAATTTGTATCCAGCACTTTGGAGGCGGAGGAGCG 13717
Qy 681 GGATCAAGAGTTAGAGATTCGAGCATCTGGCTTACACGGGTAACTGGTCTCTAC 740
Db 13716 GGATCGAGGCTGAGAGATTTGAGACCATCTGCTTACATGGTGAATCCCACTCTTAC 13657
Qy 741 TAAAAATACAAAAATTAAGCCGGGTGGTGGCAGACGCTGTATGCTCCAGCTGTCAAG 800
Db 13656 TAAAAATACAAAAATTAAGCAGGCGGTGGTGGCGGGTGGCTGTATGCTCCAGCTACTG 13597
Qy 801 AGGCTGAGGCGAGAGATGTGTGAACCGGAGGCGGAGCTTGCAGTGACCGGAGATTG 860
Db 13596 AGGCTGAGGCGAGAGATGTGTGAACCGGAGGCGGAGCTTGCAGTGAGCGGAGATTG 13537
Qy 861 CGCACTGCACTCCAGCTGGGCGAGAGTGCAGACTGTCTCAAAAAAATTTTAAAAA 920
Db 13536 CGTCACTGCTCTCAGCTGGGCGAGAGGAGGAGCTGTCTCAAAAAAATTTTAAAAA 13477
Qy 921 AAAAGTTGAATGTCTTTGAGCCCAAGTGT 950
Db 13476 GAAAAAATAAATCTCTGTAACATTAAGT 13447

RESULT 14

US-09-544-398B-11
; Sequence 11, Application US/09544398B
; Patent No. 6770461
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Recker, Robert R.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/09/544,398B
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 66933
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-544-398B-11

Query Match 5.7%; Score 287; DB 4; Length 66933;

Best Local Similarity 66.8%; Pred. No. 6.8e-52; Mismatches 230; Indels 33; Gaps 7;

Matches 528; Conservative 0;

Qy 155 CCCCAGTCCAGCTATTTGGAGGCTGAGTAGAGAGATGCTTGAAGCCAGAGATTGA 214

Db 37990 CCTTATATCCAGCATTTTGGAGTCCGAGGTGGGAGATCACTTAGAGTCAAGATTGG 38049

Qy 215 AGACAAAGCTAGGCAACATAGTGAAGCCTGTGTCTTATAAAAA---ATTAATTAGTGT 270

Db 38050 AGACCAAGCTGCGCAACATAGTGAAGCCTGTGTCTTATAAAAAATTAAGTGGG 38109

Qy 271 TGTCTTGGAGAGGCTGCA--GCTAGCTACTCGAAGAGCTGAGGTGGAGATCAC-TG 327

Db 38110 TGTGTGGCGCCGACCTGTATCCAGTTACTAGAGGCTGAGGAGAGAAATCGCTTG 38169
Qy 328 AGCCGAGAGGCTGAGGCTGCACTGAGTGAACAGTATCAACCAGTGGATTCCAGCTGGAAG 387
Db 38170 TACCCGGAGGAGAGAGTGTGAGTGAACAGATTTGTGCTGTGAGTCCAGGCTGGGTA 38229
Qy 388 ACAGAGGAGAC--CTGTCTTCCAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 445
Db 38230 ACAGAGTGAAGTCTGTGTCTCAAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 38289
Qy 446 ATAACTTGAACCTGGGACATAACTTTATGTATGATGAATTTCAATCTTTTGAAGAA 505
Db 38290 TTCTTGTATTTATATCATCAAGGCTGAGTGTCTTCCAGCTTTGTTGTCTGAAT 38349
Qy 506 TTAGCATTTCTGATATAATATATTAATTAATTAATTAATTAATTAATTAATTAATTA 554
Db 38350 AGGCGCATATGAACATTTCTTCCGTTGGAAGGTTTGTATAGTGAAGAAAGGCGTG 38409
Qy 555 -----TGAATTAATAATTTGAGAAACTAGCTTCTCACTCTCTCACTTGTCACT 604
Db 38410 TTTGTGTGTCTCATAGATTTCTGTGAAGAAATTTGCACTAACATTAATTTGTACT 38469
Qy 605 CAA--AACTTATATGTTCTTTGCGCGGCTGAGGCTCAAGCTGTATTTCCAGCACTT 662
Db 38470 ACATTAATAATGAAGCTCAGGAGGCGGCGGCTGAGTCAAGCTGTATTTCCAGCACTT 38529
Qy 663 TGGAGGCGGAGGCGGCTGATCACAAGTTAGAGATGAGAACATCTCTGCTTACACG 722
Db 38530 TGGAGGCGGAGGCGGCTGATCACAAGTTAGAGATGAGAACATCTCTGCTTACACG 38589
Qy 723 GTGAACCTGTCTCTATCAATAAATTA--CAAAAAATTTAGCGGCTGAGTCCAGAGCT 781
Db 38590 GTGAACCTGTCTCTATCAATAAATTA--CAAAAAATTTAGCGGCTGAGTCCAGAGCT 38649
Qy 782 GTAGTCCGAGCTGCTCAGAGGCTGAGGAGAGATGAGTGTGAACCCGAGAGCGGAGC 841
Db 38650 GTAGTCCGAGCTGCTCAGAGGCTGAGGAGAGATGAGTGTGAACCCGAGAGCGGAGC 38709
Qy 842 TTGAGTGAAGCCGAGATTTGCGCACTGCACTCAGCTGGGCGAGAGTCTGTGT 901
Db 38710 TTGCGGTGAGCCGAGATGCTTCACTGCACTGAGCTGGGCGAGAGACTCCGT 38769
Qy 902 CTCAAAAAAA 912
Db 38770 CTCAGCAAAA 38780

RESULT 15

US-09-543-771-11
; Sequence 11, Application US/09543771
; Patent No. 6780609
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 62
; SEQ ID NO 11
; LENGTH: 66933
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-543-771-11

Query Match 5.7%; Score 287; DB 4; Length 66933;

Best Local Similarity 66.8%; Pred. No. 6.8e-52;

Matches 528; Conservative 0; Mismatches 230; Indels 33; Gaps 7;

Qy 38110 TGTGTGGCGCCGACCTGTATCCAGTTACTAGAGGCTGAGGAGAGAAATCGCTTG 38169

Qy 328 AGCCGAGAGGCTGAGGCTGCACTGAGTGAACAGTATCAACCAGTGGATTCCAGCTGGAAG 387

Db 38170 TACCCGGAGGAGAGAGTGTGAGTGAACAGATTTGTGCTGTGAGTCCAGGCTGGGTA 38229

```
QY 155 CCCCAGTGCAGTATTGGAGGCTGAGAGAAATGCTTGAAGCCAGAGTTGA 214
Db 37990 CCTTATATCCAGATTGTTGGAGTCCGAGGTGGCAGATCAGTGTAGAGATTGG 38049
QY 215 AGACAAGCCTAGGCAATAGTGAAGCTGTGTCTATAAAAA-----ATAATTAGCTGT 270
Db 38050 AGACAGCCTGGCCAAATAGTGAACCTGGCTCTATATAAAGTACAAAAATTAGCTGGG 38109
QY 271 TGTCTTGGCACAGGCTGTCA--GCTAGCTACTGGAAAGTGAAGTGGAGATCAC-TG 327
Db 38110 TGTGTGGCCGCGACCTGTATCTCAGAGGCTGAGGCGAGAGAAATCGCTTG 38169
QY 328 AGCCAGGAGGCTGAGGCTGAGTGAACATGATCACCCAGCTGGAATCCAGCCTGGAAG 387
Db 38170 TACCCGAGAGCGAGAGGTTGAGTGAACAGAGATGTGCTGTGCAAGTCCAGCCTGGGTA 38229
QY 388 ACAGAGGAGAGAC--CCTGTTCCAAAAAATGCAAGAAAAATGCAAGAAAAAGACATC 445
Db 38230 ACAGAGTGAAGACTTCTGTCTCAAAAAAATGCAAGAAAAATCATCGAGTGAAGAGGACAC 38289
QY 446 ATAAACTTGACCTGGGACATACTTTATGTGATGAATTCACATCTTTTAGAAGAAA 505
Db 38290 TTTCTGTATTATTCATCCATCCAGGCTGAGTTCTTCCACCTTGTGTGTGTGATA 38349
QY 506 TTAGCATTTCTGATAAAATGATATTATATATTATTAATTCAA-----554
Db 38350 AGGCCACTATGAACATTTCTCCGTGTGAGATTGTTGTAAGTGAAGAAAAAGCGTG 38409
QY 555 -----TGAAATTAAATATCTGAGAACTAGCTTCTCACTCTTCAGTTGTCAGT 604
Db 38410 TTTGTGTGTGATAGGATTCTGTGTAAGAAAGTTGCACTAACATTAAGTATTGTACT 38469
QY 605 CAA--AACTTATATGCTTTTGGCCGGGTGCGGTGCTCAAGCTGTATATCCAGACTT 662
Db 38470 ACAAATAAATGAAGCTCAGAGGCGCGGCTGCTCAAGCTGTATATCCAGACTT 38529
QY 663 TGGGAGGCGCAGGCGGGTGAATCAAGATTAGAGATCGAGCCATCTGAGTAAACAG 722
Db 38530 TGGGAGGCGCAGGCGGGCGGATCATGAGTCAAGAGATCAAGACCATCTGGCCAACTG 38589
QY 723 GTGAAAACCTGTCTTACTATAAATA-CAAAAAATTAGCCGGGTGCGGTGCCAGAGCCT 781
Db 38590 GTGAAAACCCGTCTTACTATAAATAACAAAAAATTAGCCAGGTGTGTGGCGGCACT 38649
QY 782 GTAGTCCGAGCTGTCTGAGAGGCTGAGGCGAGAAATGTGTGAACCCGGAGAGCGGAGC 841
Db 38650 GTAGTCCGAGCTTCTGTGAGGCTGAGGCGAGAAATGTGTGAACCCGGAGAGCGGAGC 38709
QY 842 TTGCAGTGAGCCGAGATTGCGCACTGCACTCCAGCTGGGCGAGAGTGGAGACTCTGT 901
Db 38710 TTGCGGTGAGCCGAGATCGCTTCACTGCACTCGAGCTGGCAACAGAGCAAGATCCGT 38769
QY 902 CTCAAAAAAA 912
Db 38770 CTCACGCAAAA 38780
```

Search completed: December 27, 2004, 11:07:41
Job time : 268.615 sec8

QY 181 TGAGGTAGAGGAAATGCTTGAAGCCAGAGATTGAAGCAAGCCTTAGGCAACATATGTAGA 240
| | | | |
Db 181 TGAGGTAGAGGAAATGCTTGAAGCCAGAGATTGAAGCAAGCCTTAGGCAACATATGTAGA 240
QY 241 CCCTGTCTATAAAAAATAATATAGTGTGTCTTGGCAGAGGCTGCAGCTAGTACT 300
| | | | |
Db 241 CCCTGTCTATAAAAAATAATATAGTGTGTGTCTTGGCAGAGGCTGCAGCTAGTACT 300
QY 301 CGGAAAGCTGAGGTGGAGGATCACTAGAGCCAGAGAGGCTGAGGCTGCAAGTGA 360
| | | | |
Db 301 CGGAAAGCTGAGGTGGAGGATCACTAGAGCCAGAGAGGCTGAGGCTGCAAGTGA 360
QY 361 TCACCCAGCTGAGATTCAGAGCTGGAAGACAGAGGAGACCCCTGTTCCAAAAA 420
| | | | |
Db 361 TCACCCAGCTGAGATTCAGAGCTGGAAGACAGAGGAGACCCCTGTTCCAAAAA 420
QY 421 AAAAAAATGCAAGAAAGACATCATAACTTGACCTGGGACATACTTTATGTATG 480
| | | | |
Db 421 AAAAAAATGCAAGAAAGACATCATAACTTGACCTGGGACATACTTTATGTATG 480
QY 481 AAATTCATCTTTTATGAAAGAAATTAGCAATTTTGATTAATATTAATTAAT 540
| | | | |
Db 481 AAATTCATCTTTTATGAAAGAAATTAGCAATTTTGATTAATATTAATTAAT 540
QY 541 ATTATAAATCAATGGAATTAATATTTCTGAGAACTAGCTCTCACTCTCAAGTGT 600
| | | | |
Db 541 ATTATAAATCAATGGAATTAATATTTCTGAGAACTAGCTCTCTCAAGTGT 600
QY 601 CAGTCAAAATTTAATAGTCTTTGGCCGGGTGGCTCAGGCTGTAAATCCAGCAC 660
| | | | |
Db 601 CAGTCAAAATTTAATAGTCTTTGGCCGGGTGGCTCAGGCTGTAAATCCAGCAC 660
QY 661 TTTGGAGGCGGAGGCGGTGTGATCAGAGATTGAGATCGAGACCATCTGGCTTACA 720
| | | | |
Db 661 TTTGGAGGCGGAGGCGGTGTGATCAGAGATTGAGATCGAGACCATCTGGCTTACA 720
QY 721 CGGTGAAACCTGCTCTCTAATAAAAAATTAAGCCGGGTGGGTGCGGAGCGCC 780
| | | | |
Db 721 CGGTGAAACCTGCTCTCTAATAAAAAATTAAGCCGGGTGGGTGCGGAGCGCC 780
QY 781 TGTATGTCAGCTGTCTCAGAGGCTGAGGAGAAATGTGTGAACCCGGAGGCGAG 840
| | | | |
Db 781 TGTATGTCAGCTGTCTCAGAGGCTGAGGAGAAATGTGTGAACCCGGAGGCGAG 840
QY 841 CTTCGAGTGGCGGAGATTGGCCACTGCACTCAGCTGGGAGACAGTGGCGAGACTCG 900
| | | | |
Db 841 CTTCGAGTGGCGGAGATTGGCCACTGCACTCAGCTGGGAGACAGTGGCGAGACTCG 900
QY 901 TCTCAAAAAAAGTTGAATGCTTTTGAAGCCAGTATGCTTCTTTTC 960
| | | | |
Db 901 TCTCAAAAAAAGTTGAATGCTTTTGAAGCCAGTATGCTTCTTTTC 960
QY 961 TTTCTTCTCTTTTCTTTTCTTCAAAAAATCTCTAGATTGAATCTTGAATTTGGCT 1020
| | | | |
Db 961 TTTCTTCTCTTTTCTTTTCTTCAAAAAATCTCTAGATTGAATCTTGAATTTGGCT 1020
QY 1021 TAAGTCTCTCTCTGAGCAATTTTGAATGAATAATACAGTCAATTAATTAATTC 1080
| | | | |
Db 1021 TAAGTCTCTCTCTGAGCAATTTTGAATGAATAATACAGTCAATTAATTAATTC 1080
QY 1081 CTGAAATTTTAAAAAATCATGAGGTTCAAAATTCATTAATTTTGTGT 1140
| | | | |
Db 1081 CTGAAATTTTAAAAAATCATGAGGTTCAAAATTCATTAATTTTGTGT 1140
QY 1141 GATTAATGACATACCTTAATTTTCCCTTAATTAATGATTTTAATTAATCTCATTC 1200
| | | | |
Db 1141 GATTAATGACATACCTTAATTTTCCCTTAATTAATGATTTTAATTAATCTCATTC 1200
QY 1201 TGTCTCACTTATGATTAATTAATTAATGAGCAATTAATTAAGAAATCAACAGGAA 1260
| | | | |
Db 1201 TGTCTCACTTATGATTAATTAATTAATGAGCAATTAATTAAGAAATCAACAGGAA 1260

QY 1261 AGTGTGCAATTTGAAGACTTACAGACTGCAAAACCAATTTTAAGCCAGATTCCTTGACATGTT 1320
| | | | |
Db 1261 AGTGTGCAATTTGAAGACTTACAGACTGCAAAACCAATTTTAAGCCAGATTCCTTGACATGTT 1320
QY 1321 GTGCTGTATATATAGTACTTTTCAATATAGTAAACATTAATTAATTAATGATGGAAGAG 1380
| | | | |
Db 1321 GTGCTGTATATATAGTACTTTTCAATATAGTAAACATTAATTAATTAATGATGGAAGAG 1380
QY 1381 CAAGCAAGAAAGAAAGTATTTCAATCAATCTCTCTCTCCATCAGCATTTGGCTTA 1440
| | | | |
Db 1381 CAAGCAAGAAAGAAAGTATTTCAATCAATCTCTCTCTCCATCAGCATTTGGCTTA 1440
QY 1441 ATATCATCAATTTGTCAGTTAAGAAACAATAGGCTTACCAATAGTTTTTGAATTA 1500
| | | | |
Db 1441 ATATCATCAATTTGTCAGTTAAGAAACAATAGGCTTACCAATAGTTTTTGAATTA 1500
QY 1501 TGAATGAATGGCAACCCCTCTAAGACTATTTGATACATTAATTTTGAAGCAAGAGAT 1560
| | | | |
Db 1501 TGAATGAATGGCAACCCCTCTAAGACTATTTGATACATTAATTTTGAAGCAAGAGAT 1560
QY 1561 GCAATGATATTTTCAACTTTTTCCTGTTTATGATTTCTGTGTTCTTTGACTATA 1620
| | | | |
Db 1561 GCAATGATATTTTCAACTTTTTCCTGTTTATGATTTCTGTGTTCTTTGACTATA 1620
QY 1621 AAGTTAGCTAGGTAAGCAATTTGTTTAAAGTCTGAAAAACAAATGCTTTCAGATTA 1680
| | | | |
Db 1621 AAGTTAGCTAGGTAAGCAATTTGTTTAAAGTCTGAAAAACAAATGCTTTCAGATTA 1680
QY 1681 GGTAGGAGAAAAATATCTCTCAACATGTCACATTTAGCAGGAAAACTTAATATCA 1740
| | | | |
Db 1681 GGTAGGAGAAAAATATCTCTCAACATGTCACATTTAGCAGGAAAACTTAATATCA 1740
QY 1741 TATCACATCAATGATATATTAATTAATATCATTGATATAGTAAAGTCAATCTTA 1800
| | | | |
Db 1741 TATCACATCAATGATATATTAATTAATATCATTGATATAGTAAAGTCAATCTTA 1800
QY 1801 AAATATGATATCAATATAGCACTTAACCTTGTGGCAGAAACAAACCTTAACGTGCA 1860
| | | | |
Db 1801 AAATATGATATCAATATAGCACTTAACCTTGTGGCAGAAACAAACCTTAACGTGCA 1860
QY 1861 TTTTATCTAATCAATTAACAGCTGCTCGTTTTCAGTTGTGCAATCGAATGCAAGCA 1920
| | | | |
Db 1861 TTTTATCTAATCAATTAACAGCTGCTCGTTTTCAGTTGTGCAATCGAATGCAAGCA 1920
QY 1921 TCCCTGTCTGATGAGGATTTTCTGCACTGATTAAGGAAAAATGCTGAAAGTTGTAGGCT 1980
| | | | |
Db 1921 TCCCTGTCTGATGAGGATTTTCTGCACTGATTAAGGAAAAATGCTGAAAGTTGTAGGCT 1980
QY 1981 GCTCAGAGCAGACATCATGATGATCATATGAAGCTCAAGCTGCTGACCTTGGCAA 2040
| | | | |
Db 1981 GCTCAGAGCAGACATCATGATGATCATATGAAGCTCAAGCTGCTGACCTTGGCAA 2040
QY 2041 AAAGGAGAGAAAGAGATGAGAGAGCAGTGGGAGAAAGTTCAAGTGGGTTTTCTC 2100
| | | | |
Db 2041 AAAGGAGAGAAAGAGATGAGAGAGCAGTGGGAGAAAGTTCAAGTGGGTTTTCTC 2100
QY 2101 CTGGAACCTCAAGATTAATAGGCTCAAGAGCTGTGTGCAAAACCTTAACAGCTGATCA 2160
| | | | |
Db 2101 CTGGAACCTCAAGATTAATAGGCTCAAGAGCTGTGTGCAAAACCTTAACAGCTGATCA 2160
QY 2161 GCTGCTACATTTGTCAGAGGAGGCGAGGACACAGAGTGTGAGAGAGTCAAGCTGAGA 2220
| | | | |
Db 2161 GCTGCTACATTTGTCAGAGGAGGCGAGGACACAGAGTGTGAGAGAGTCAAGCTGAGA 2220
QY 2221 CCACGAGAGGCTGCGTTCAAGGATTTTGTATCCAGAGAGAGCATCTTCTCTATGATA 2280
| | | | |
Db 2221 CCACGAGAGGCTGCGTTCAAGGATTTTGTATCCAGAGAGAGCATCTTCTCTATGATA 2280
QY 2281 AACCAAGAGATTCAGACCTCTTTTGTAGCGGAGCTGATCTTCTGCGGATAGTCT 2340
| | | | |
Db 2281 AACCAAGAGATTCAGACCTCTTTTGTAGCGGAGCTGATCTTCTGCGGATAGTCT 2340
QY 2341 AAACCAATTAATGAATAATCTATTAAGTCAAGAAAAATTAATGCTGTATGATCAAA 2400
| | | | |

QY 4561 TGCATTAAGAGCACTTTTATTTATTTCTAAGCTCTCTCATATTCATTTGGCCATGGT 4620
Db 4561 TGCATTAAGAGCACATTTTATTTATTTCTAAGCTCTCTCATATTCATTTGGCCATGGT 4620
QY 4621 AACTGTTTAAAGGTGGCTGGGAACGGGGGACCCCTGGAACTATCTTGATTTACATGGGCA 4680
Db 4621 AACTGTTTAAAGGTGGCTGGGAACGGGGGACCCCTGGAACTATCTTGATTTACATGGGCA 4680
QY 4681 CCATGACACTTCTGATCTCTCTCTCTGAGTCTGATCTTGATTTGTTCTGACAGACCTT 4740
Db 4681 CCATGACACTTCTGATCTCTCTCTCTGAGTCTGATCTTGATTTGTTCTGACAGACCTT 4740
QY 4741 TCCAGCCCGAGTTTAAAGAAATTCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4800
Db 4741 TCCAGCCCGAGTTTAAAGAAATTCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4800
QY 4801 TTTTCATTTAATCT 4860
Db 4801 TTTTCATTTAATCT 4860
QY 4861 TCCATTTATTAAGATTTCT 4920
Db 4861 TCCATTTATTAAGATTTCT 4920
QY 4921 ATCTCAAGAAATCTGAAATCT 4980
Db 4921 ATCTCAAGAAATCTGAAATCT 4980
QY 4981 CTATCTTTCTTATATTTGT 5000
Db 4981 CTATCTTTCTTATATTTGT 5000

RESULT 2
US-10-224-562-3
Sequence 3, Application US/10224562
Publication No. US200302229A1
GENERAL INFORMATION:
APPLICANT: YAN, Chunhua et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001098DIV
CURRENT APPLICATION NUMBER: US/10/224,562
CURRENT FILING DATE: 2002-08-21
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 53332
TYPE: DNA
ORGANISM: Homo sapiens
US-10-224-562-3

Query Match 100.0%; Score 5000; DB 14; Length 53332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TATAGGCCAATGCTGTGGCTCAAGCGCTGATATTCACGACACTTTGGAGGACAGAGATCG 60
Db 1 TATAGGCCAATGCTGTGGCTCAAGCGCTGATATTCACGACACTTTGGAGGACAGAGATCG 60
QY 61 CTGAGCTCAGAAATTGGAGACAAAGCTACGTAACTAGTGAAGAACTCTGTCTGTACAAA 120
Db 61 CTGAGCTCAGAAATTGGAGACAAAGCTACGTAACTAGTGAAGAACTCTGTCTGTACAAA 120
QY 121 TATATAAAGAAATTTTCAGGCAATGTTGGCGTGCAACCCCAAGTGCAGCTATTTGGAGGC 180
Db 121 TATATAAAGAAATTTTCAGGCAATGTTGGCGTGCAACCCCAAGTGCAGCTATTTGGAGGC 180
QY 181 TGAAGTGAAGAGAAAGCTTGAAGCCGAGGAGTTGAAGACAAAGCTTAGGCAACATAGTGAGA 240
Db 181 TGAAGTGAAGAGAAAGCTTGAAGCCGAGGAGTTGAAGACAAAGCTTAGGCAACATAGTGAGA 240

QY 241 CCTGTGTCTATATAAAAAATTAATAGCTGGTGTCTTGGACAGGCTCGACGTAGTACT 300
Db 241 CCTGTGTCTATATAAAAAATTAATAGCTGGTGTCTTGGACAGGCTCGACGTAGTACT 300
QY 301 CGGAATACAGAGTGGGAGAGATTCATGAGCCAGAGAGCTGAGGCTGACGTAAACGTGA 360
Db 301 CGGAATACAGAGTGGGAGAGATTCATGAGCCAGAGAGCTGAGGCTGACGTAAACGTGA 360
QY 361 TCACCCAGCTGATTCACAGCTGGAAGACAGAGGAGACCTCTGTTTCAAAAAA 420
Db 361 TCACCCAGCTGATTCACAGCTGGAAGACAGAGGAGACCTCTGTTTCAAAAAA 420
QY 421 AAAAAAATGCAAGAAAGACATCAATTAATCTGAGCTGGAACATTAATTTATGTGATG 480
Db 421 AAAAAAATGCAAGAAAGACATCAATTAATCTGAGCTGGAACATTAATTTATGTGATG 480
QY 481 AAATTCACAACTTTTATAGAAAGAAATTTGCAATTTCTGATTAATATTAATATTT 540
Db 481 AAATTCACAACTTTTATAGAAAGAAATTTGCAATTTCTGATTAATATTAATATTT 540
QY 541 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
Db 541 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
QY 601 CAGTCAAAATCTTAAATGCTCTTGGCCGGGTGCGGTGCTCAGCCTGATATCCAGCAC 660
Db 601 CAGTCAAAATCTTAAATGCTCTTGGCCGGGTGCGGTGCTCAGCCTGATATCCAGCAC 660
QY 661 TTTGGAGGCGAGGCGGGTGGATCAAGATTTAGAGATGAGACATCTCGGCTAACCA 720
Db 661 TTTGGAGGCGAGGCGGGTGGATCAAGATTTAGAGATGAGACATCTCGGCTAACCA 720
QY 721 CGGTGAACCTGCT 780
Db 721 CGGTGAACCTGCT 780
QY 781 TGTATCCCAAGTGTCTCAAGAGGCTGAGAGCAGAGAAATGTTGTGAACCCGGAGGCGAG 840
Db 781 TGTATCCCAAGTGTCTCAAGAGGCTGAGAGCAGAGAAATGTTGTGAACCCGGAGGCGAG 840
QY 841 CTTCAGTGAAGCGAGATTTGGCCACTGCACTCCAGCCTGGGCGACAGTGCAGACTGTG 900
Db 841 CTTCAGTGAAGCGAGATTTGGCCACTGCACTCCAGCCTGGGCGACAGTGCAGACTGTG 900
QY 901 TCTCAAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAA 960
Db 901 TCTCAAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAA 960
QY 961 TTTCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1020
Db 961 TTTCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1020
QY 1021 TAAAGTCTCTCTCTCTGAGCAATTTTGAATGAAGAAATTAATCATGCTCTAAATTAAT 1080
Db 1021 TAAAGTCTCTCTCTCTGAGCAATTTTGAATGAAGAAATTAATCATGCTCTAAATTAAT 1080
QY 1081 CTGAACATTTTAAAAAATCATCATGAGGTTCAAAATATCAATTAATTAATTTGTTGT 1140
Db 1081 CTGAACATTTTAAAAAATCATCATGAGGTTCAAAATATCAATTAATTAATTTGTTGT 1140
QY 1141 GATATTAAGCAATCTCTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1200
Db 1141 GATATTAAGCAATCTCTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1200
QY 1201 TGTCTCACTTTATGATTTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
Db 1201 TGTCTCACTTTATGATTTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
QY 1261 AGTGTGATTTGAAGACTACAGACTGCAACCAATTTAAGCCAGATTCCTTGAACATGTT 1320
Db 1261 AGTGTGATTTGAAGACTACAGACTGCAACCAATTTAAGCCAGATTCCTTGAACATGTT 1320

QY 1321 GNGCTGTAATAGTACTTTACATATAGTAACTTAACTTATACATATATGTGAGAGAG 1380
DB 1321 GNGCTGTAATAGTACTTTACATATAGTAACTTAACTTATACATATATGTGAGAGAG 1380
QY 1381 CAAGCAAGAAAGAGAAAGATATTTCTTCAAACTCTCTCTCCATCAGCATTTGGCTA 1440
DB 1381 CAAGCAAGAAAGAGAAAGATATTTCTTCAAACTCTCTCTCCATCAGCATTTGGCTA 1440
QY 1441 ATATCATCTTTGTACAGTTAAGAACATAGGTGTCTCACATATGTTTTGAATAAA 1500
DB 1441 ATATCATCTTTGTACAGTTAAGAACATAGGTGTCTCACATATGTTTTGAATAAA 1500
QY 1501 TGAATGAAGGCAACCTTCTAGACATATGGAATACATCTTGTGTAAGGCAAGAGAT 1560
DB 1501 TGAATGAAGGCAACCTTCTAGACATATGGAATACATCTTGTGTAAGGCAAGAGAT 1560
QY 1561 GCAGTAGATATTTTCAACTTTTCTGTTTATGATCTGTGGTTCTTTGACTACTAA 1620
DB 1561 GCAGTAGATATTTTCAACTTTTCTGTTTATGATCTGTGGTTCTTTGACTACTAA 1620
QY 1621 AAGTTAGCTAGTAGCAAAATTTGTTTAAAGTCTGAAAACCAAAATGCTTTCAGATPAAA 1680
DB 1621 AAGTTAGCTAGTAGCAAAATTTGTTTAAAGTCTGAAAACCAAAATGCTTTCAGATPAAA 1680
QY 1681 GGTAGGAGAAAATCTCTCAACATGTCCACTTTAGCACAGAAAACCTTAATATCAA 1740
DB 1681 GGTAGGAGAAAATCTCTCAACATGTCCACTTTAGCACAGAAAACCTTAATATCAA 1740
QY 1741 TATCACCATCATGATATCATATATATATCATTTGATGATAGATGATGATGATGATGAT 1800
DB 1741 TATCACCATCATGATATCATATATATATCATTTGATGATAGATGATGATGATGATGAT 1800
QY 1801 AAACTATGATATCAATAGCACTACTGTGGCCAGAACAAAGCCTTAACTGTGCCAAA 1860
DB 1801 AAACTATGATATCAATAGCACTACTGTGGCCAGAACAAAGCCTTAACTGTGCCAAA 1860
QY 1861 TTTTATTTTATTAATTAACAGCTGCTCTGTTTCACTGTGCACTGTGAAATGCAAGCAA 1920
DB 1861 TTTTATTTTATTAATTAACAGCTGCTCTGTTTCACTGTGCACTGTGAAATGCAAGCAA 1920
QY 1921 TCCCTCTGATGTGAGATTTCTTGCACTGATAGGAAAACCTGTAAGTTGAGAGCT 1980
DB 1921 TCCCTCTGATGTGAGATTTCTTGCACTGATAGGAAAACCTGTAAGTTGAGAGCT 1980
QY 1981 GCTCCAGGAGAGCACTCATGTGATCATATGAAAGCTCCAGCTGTGCACTGTGCA 2040
DB 1981 GCTCCAGGAGAGCACTCATGTGATCATATGAAAGCTCCAGCTGTGCACTGTGCA 2040
QY 2041 AAAGGAGAGAACATAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
DB 2041 AAAGGAGAGAACATAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
QY 2101 CTTGAACCTTACAGATTTATGAGGTCAAGAGCTGTGCAAAAGCTGTACAGCTGTGAGCA 2160
DB 2101 CTTGAACCTTACAGATTTATGAGGTCAAGAGCTGTGCAAAAGCTGTACAGCTGTGAGCA 2160
QY 2161 GCTGCTACCATTTGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
DB 2161 GCTGCTACCATTTGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
QY 2221 CCAGGAGAGCTGTGTTCAAGATTTTGTATCCAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
DB 2221 CCAGGAGAGCTGTGTTCAAGATTTTGTATCCAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
QY 2281 AACCAAGAGATTTAGACACTCCCTTTTGTGAGGAGATGATTTCTTCTGAGTAAAGTCT 2340
DB 2281 AACCAAGAGATTTAGACACTCCCTTTTGTGAGGAGATGATTTCTTCTGAGTAAAGTCT 2340
QY 2341 AAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2400
DB 2341 AAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2400
QY 2401 TTTGGGGAATTTCTTGTAAACCAAAAGGAGAAAATAATCTTGGCTTGGGCTGACGAA 2460

DB 2401 TTTGGGGAATTTCTTGTAAACCAAAAGGAGAAAATAATCTTGGCTTGGGCTGACGAA 2460
QY 2461 ACTCACTGTGCTTGAAGTGCAGAAAGATAGTCTCTCAAAATCTTAAGGTCTTAAATTA 2520
DB 2461 ACTCACTGTGCTTGAAGTGCAGAAAGATAGTCTCTCAAAATCTTAAGGTCTTAAATTA 2520
QY 2521 AAGCTGAAAATTAAAGGCAAGCTGCAATATTAGTTGATGCTATGATTTGAAACTT 2580
DB 2521 AAGCTGAAAATTAAAGGCAAGCTGCAATATTAGTTGATGCTATGATTTGAAACTT 2580
QY 2581 TGGTAAATTTGTCATGATTTATAGCAATGCAATGATTTTCCCTACAGCAATTAATTA 2640
DB 2581 TGGTAAATTTGTCATGATTTATAGCAATGCAATGATTTTCCCTACAGCAATTAATTA 2640
QY 2641 AGTGAACATGAATAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700
DB 2641 AGTGAACATGAATAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700
QY 2701 AGAGATTTAGTTAACTGATGTTAAAGACATTTTAAAGTGAAGAAAGATGATGA 2760
DB 2701 AGAGATTTAGTTAACTGATGTTAAAGACATTTTAAAGTGAAGAAAGATGATGA 2760
QY 2761 TTTTCAAGAGTTGATACATTTATGATGCTTTTCAAGGAGATCTTCAAGAAAAGTGCCTTTG 2820
DB 2761 TTTTCAAGAGTTGATACATTTATGATGCTTTTCAAGGAGATCTTCAAGAAAAGTGCCTTTG 2820
QY 2821 GGGGTACAGAAAGCTTGAAGAAACATTTGAAGAGTGAAGAAATGAGCAATTAAGAAAAT 2880
DB 2821 GGGGTACAGAAAGCTTGAAGAAACATTTGAAGAGTGAAGAAATGAGCAATTAAGAAAAT 2880
QY 2881 GGTTTTACAGAGCACTGAATCTTTTACTTTGATTAATTTTATTTTCTGCTTTCTTTT 2940
DB 2881 GGTTTTACAGAGCACTGAATCTTTTACTTTGATTAATTTTATTTTCTGCTTTCTTTT 2940
QY 2941 CTCTAGCTAACAGACTTAAAGAGAGATGATGATGATGATGATGATGATGATGATGATG 3000
DB 2941 CTCTAGCTAACAGACTTAAAGAGAGATGATGATGATGATGATGATGATGATGATGATG 3000
QY 3001 CTTCAAGCTGCCGCTGCCAGAAAGTTCAAGAGTAAAGGCAAGAGTAACTGATGATG 3060
DB 3001 CTTCAAGCTGCCGCTGCCAGAAAGTTCAAGAGTAAAGGCAAGAGTAACTGATGATG 3060
QY 3061 TTTTCAAGAGAGAGATGAGGAGAGGTTTCAAGTGAAGTGAAGTGAAGTGAAGTGA 3120
DB 3061 TTTTCAAGAGAGAGATGAGGAGAGGTTTCAAGTGAAGTGAAGTGAAGTGAAGTGA 3120
QY 3121 AAGAGAAATTTATATGAGACTTGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180
DB 3121 AAGAGAAATTTATATGAGACTTGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180
QY 3181 GAGCGAGCTTCCCAAGTCTGCTGAGCAATGCTGTCTAATTTTCCCTGGGAGAAAAGT 3240
DB 3181 GAGCGAGCTTCCCAAGTCTGCTGAGCAATGCTGTCTAATTTTCCCTGGGAGAAAAGT 3240
QY 3241 CAACTAATAAATAAAGTCTTTCTCTCTCTCTTCAACCGCTCTTTTCCCAATTC 3300
DB 3241 CAACTAATAAATAAAGTCTTTCTCTCTCTCTTCAACCGCTCTTTTCCCAATTC 3300
QY 3301 CCTAGAGAGAGAGAGAGAGCTCTTTTGGGAGAGCTCATCTTAACTTGAAGTGAAG 3360
DB 3301 CCTAGAGAGAGAGAGAGAGCTCTTTTGGGAGAGCTCATCTTAACTTGAAGTGAAG 3360
QY 3361 GCTGGGTGAAGCTCTTATGAGCAAGTTTAAAGGAGATTAAGAGTGAAGTGAAGTGA 3420
DB 3361 GCTGGGTGAAGCTCTTATGAGCAAGTTTAAAGGAGATTAAGAGTGAAGTGAAGTGA 3420
QY 3421 CTGGGAGAGACTTTAGATGAGAGTCCGCCCCCAATTTTCAATTTTAAAGCCAGGT 3480
DB 3421 CTGGGAGAGACTTTAGATGAGAGTCCGCCCCCAATTTTCAATTTTAAAGCCAGGT 3480
QY 3481 GAGACATCATTAAGTCTTACACTCAGAGCTGTGCAAGACATAGGCGGAGAGG 3540

Db 3481 GAGACATCATAGAAATTCATAGCACTGAGAACCTGTGCAAGACCAATGGCCGACAGGGA 3540
QY 3541 GAGAGACATGATTAACCTTAACACGCTTGAAGAAAAAACAACCTGCCCTGCTTAATTA 3600
Db 3541 GAGAGACATGATTAACCTTAACACGCTTGAAGAAAAAACAACCTGCCCTGCTTAATTA 3600
QY 3601 AATGAGCCCACTTAAGTTTATGACGCTTCCCTTCTTGATTCATATTCAGAGAAATTC 3660
Db 3601 AATGAGCCCACTTAAGTTTATGACGCTTCCCTTCTTGATTCATATTCAGAGAAATTC 3660
QY 3661 AAGAAAAATGACATTCCTCTACTAGCAACCAAGAAACAAATATCACTTCAGGCTGTG 3720
Db 3661 AAGAAAAATGACATTCCTCTACTAGCAACCAAGAAACAAATATCACTTCAGGCTGTG 3720
QY 3721 GGAGGCACTGTTGTTAAAGGCTCTCTACAGGTTTTTTATATCCCTCTTAATCAAT 3780
Db 3721 GGAGGCACTGTTGTTAAAGGCTCTCTACAGGTTTTTTATATCCCTCTTAATCAAT 3780
QY 3781 GACAGATTTTGTAAATGGCAACCTGGAATTTGCTGCTTCAATTCCTGCACTGCTTAT 3840
Db 3781 GACAGATTTTGTAAATGGCAACCTGGAATTTGCTGCTTCAATTCCTGCACTGCTTAT 3840
QY 3841 AGAAGAACTGAAGTTGTTTGTGCAAAATATGATGATGCAAAAGATGATTAATCCTAG 3900
Db 3841 AGAAGAACTGAAGTTGTTTGTGCAAAATATGATGATGCAAAAGATGATTAATCCTAG 3900
QY 3901 AATTTTATATTTGCAAAAATACCAAAATGCTGAGAAATTAATATCTGCTTATCCAA 3960
Db 3901 AATTTTATATTTGCAAAAATACCAAAATGCTGAGAAATTAATATCTGCTTATCCAA 3960
QY 3961 AGCTAAGTACTAATTTTGTAAACCAACCACTTGTAAATATATGTAAGAAATCCATGA 4020
Db 3961 AGCTAAGTACTAATTTTGTAAACCAACCACTTGTAAATATATGTAAGAAATCCATGA 4020
QY 4021 AATCCCTTTTAACTCAAGTGGGAAAGTGGATGCTTTTCTTTATATGATCCCA 4080
Db 4021 AATCCCTTTTAACTCAAGTGGGAAAGTGGATGCTTTTCTTTATATGATCCCA 4080
QY 4081 ATGAGAGAAAGTAATAGGCTCAATAGTGGTAAATTTTAAATTAATATATGCTGAT 4140
Db 4081 ATGAGAGAAAGTAATAGGCTCAATAGTGGTAAATTTTAAATTAATATATGCTGAT 4140
QY 4141 CCGAGTGCAGTGGTGTAACTACTGATCAACAACCACTGACAGATTTCTTTGTTCC 4200
Db 4141 CCGAGTGCAGTGGTGTAACTACTGATCAACAACCACTGACAGATTTCTTTGTTCC 4200
QY 4201 TCTCCACTCCCACTGCTTCACTTACTGSCCAAAAAAGAAAAATTTTATATTA 4260
Db 4201 TCTCCACTCCCACTGCTTCACTTACTGSCCAAAAAAGAAAAATTTTATATTA 4260
QY 4261 CTACTCAAGACTAATATTTATTTATTTATCTAGATTTATGCTGTAATTTATTTT 4320
Db 4261 CTACTCAAGACTAATATTTATTTATTTATCTAGATTTATGCTGTAATTTATTTT 4320
QY 4321 ACTTGTAAAAACAGATTTGAGGGAATACAGTTTAAATTTTAAATTAATATAT 4380
Db 4321 ACTTGTAAAAACAGATTTGAGGGAATACAGTTTAAATTTTAAATTAATATAT 4380
QY 4381 TTAATTTATTTTGAATGAAATCTGTACCAACCGTGAAGTGCAGTGGCGATC 4440
Db 4381 TTAATTTATTTTGAATGAAATCTGTACCAACCGTGAAGTGCAGTGGCGATC 4440
QY 4441 TCAGATGACTGCAACCTGCTGCTGAGTTCAAGCAACCTGSCCTGSCCTTTA 4500
Db 4441 TCAGATGACTGCAACCTGCTGCTGAGTTCAAGCAACCTGSCCTGSCCTTTA 4500
QY 4501 TACTTTCTTAATCTGTTTATGTCATGATGATCCTTAATCTTTTCAATGCTGAAGATC 4560
Db 4501 TACTTTCTTAATCTGTTTATGTCATGATGATCCTTAATCTTTTCAATGCTGAAGATC 4560
QY 4561 TCGAATTAAGACCACTTTATTTATTTATTTATGCTTCTCATATCAATTTGGCCATG 4620
Db 4561 TCGAATTAAGACCACTTTATTTATTTATTTATGCTTCTCATATCAATTTGGCCATG 4620

QY 4621 AACTGTTTCAAGTGGCTGCAACGCGGCAACCTGGAACAATCTTGATATCATGGCA 4680
Db 4621 AACTGTTTCAAGTGGCTGCAACGCGGCAACCTGGAACAATCTTGATATCATGGCA 4680
QY 4681 CCATGCACTTGTATCTCTCTTCTTGAGTTCTGACTTTGATTTGTTGACAGACCTT 4740
Db 4681 CCATGCACTTGTATCTCTCTTCTTGAGTTCTGACTTTGATTTGTTGACAGACCTT 4740
QY 4741 TCCAGCCCAAGTTTACAGAAATTCATATCTTTCTTCTAGTAACTTAATGTTTTCT 4800
Db 4741 TCCAGCCCAAGTTTACAGAAATTCATATCTTTCTTCTAGTAACTTAATGTTTTCT 4800
QY 4801 TTTTCATTTACTCTTTCATCTACTGGAATTAATGTAATTTCAATCAATCCAC 4860
Db 4801 TTTTCATTTACTCTTTCATCTACTGGAATTAATGTAATTTCAATCAATCCAC 4860
QY 4861 TCCATTTATTAATTTTCTTCTCTGATGTTGAAATGCTGCCATGATTAATTA 4920
Db 4861 TCCATTTATTAATTTTCTTCTCTGATGTTGAAATGCTGCCATGATTAATTA 4920
QY 4921 ATCTCAGAAATCTTGAAATCTTCTCTTCTTAATCTTTTAAATATCATGTTCCCTAAT 4980
Db 4921 ATCTCAGAAATCTTGAAATCTTCTCTTCTTAATCTTTTAAATATCATGTTCCCTAAT 4980
QY 4981 CTATCTTTCTTAATTTGT 5000
Db 4981 CTATCTTTCTTAATTTGT 5000

RESULT 3
US-10-027-632-132158
; Sequence 132158, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108627.129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198, 676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193, 483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185, 218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167, 363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156, 358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146, 002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132158
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-132158

Query Match 13.7%; Score 607.4; DB 13; Length 714;
Best Local Similarity 99.3%; Pred. No. 2.9e-128;
Matches 710; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

QY 795 CTCAGAGGCTGAGGCAAGATGTTGAACCCGAGAGGCGAGCTTGAGTGAAGCG 854
Db 1 CTCAGAGGCTGAGGCAAGATGTTGAACCCGAGAGGCGAGCTTGAGTGAAGCG 60
QY 855 AGATTGCCCACTGCACTCCAGCCTGGGCGACAGTGCAGACTGTCTCAAAAAAAA 914
Db 855 AGATTGCCCACTGCACTCCAGCCTGGGCGACAGTGCAGACTGTCTCAAAAAAAA 914

Db 61 AGATTGGCGTCTGCACTGCGGCGGCAAGTGGGAGACTCTGTCTCAAAAAA 120
Qy 915 AAAAAAAAA-AGTGAATGCTTTTGGAGCAAGTACTCTTCTTCTTCTTCTT 973
Db 121 AAAAAAAAAAGTGAATGCTTTTGGAGCAAGTACTCTTCTTCTTCTTCTT 179
Qy 974 TTTTCTTTTCAAAAAATCTCTAGATTGAATCTTGGAAATGGCTTAAGTCTTCTC 1033
Db 180 TTTTCTTTTCAAAAAATCTCTAGATTGAATCTTGGAAATGGCTTAAGTCTTCTC 239
Qy 1034 TTGTGGCAATTTTGAATGAAAAATACATGTCATTAATTAATCTGAACTTTTAA 1093
Db 240 TTGTGGCAATTTTGAATGAAAAATACATGTCATTAATTAATCTGAACTTTTAA 299
Qy 1094 AAAACCATCATGAGGTTCAAAATATCAAAATATTAATTTGTGTGATATAGACATA 1153
Db 300 AAAACCATCATGAGGTTCAAAATATCAAAATATTAATTTGTGTGATATAGACATA 359
Qy 1154 ACTCTTAATTTTCCCTTAATAATGATTTGTTATATATCTTCAATTCCTCTCATTAT 1213
Db 360 ACTCTTAATTTTCCCTTAATAATGATTTGTTATATATCTTCAATTCCTCTCATTAT 419
Qy 1214 GATTGATATATATAGTGGCAATATCTTAGGAATCTAAAGAGAAAGTTGCAATTG 1273
Db 420 GATTGATATATATAGTGGCAATATCTTAGGAATCTAAAGAGAAAGTTGCAATTG 479
Qy 1274 AAGACTACAGACTGCAAAACCAATTTAAGCCAGATTCCTTGACATGTTGCTTTAATAT 1333
Db 480 AAGACTACAGACTGCAAAACCAATTTAAGCCAGATTCCTTGACATGTTGCTTTAATAT 539
Qy 1334 AGTACTTTACATATAGTAATAACATTATATCATATATGTGGAAGAGCAAGCAAGAAAG 1393
Db 540 AGTACTTTACATATAGTAATAACATTATATCATATATGTGGAAGAGCAAGCAAGAAAG 599
Qy 1394 AAGAAAGTATTTTCATTCAAATCTCTCTCATCAATGCGTAATATCATCATTTG 1453
Db 600 AAGAAAGTATTTTCATTCAAATCTCTCTCATCAATGCGTAATATCATCATTTG 659
Qy 1454 TACAGTTAAGAACCAATAGTGCTGCACACATAGTTTGTGAATTAATGAATGAA 1508
Db 660 GACAGTTAAGAACCAATAGTGCTGCACACATAGTTTGTGAATTAATGAATGAA 714

RESULT 4
US-10-027-632-132158
; Sequence 132158, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FaSeq for Windows Version 4.0
; SEQ ID NO 132158
; LENGTH: 714
; TYPE: DNA

ORGANISM: Human
US-10-027-632-132158
Query Match 13.7%; Score 687.4; DB 15; Length 714;
Best Local Similarity 99.3%; Pred. No. 2,9e-128;
Matches 710; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

Qy 795 CTCAGAGGCTGAGGAGGAGGAATGCTGAACCCGGAGGCGGAGCTTGAGTGGCG 854
Db 1 CTCAGAGGCTGAGGAGGAGGAATGCTGAACCCGGAGGCGGAGCTTGAGTGGCG 60
Qy 855 AGATTGGCGTCTGCACTGCGGCGGAGAGTGGCGAGCTCTGTCAAAAAA 914
Db 61 AGATTGGCGTCTGCACTGCGGCGGAGAGTGGCGAGCTCTGTCAAAAAA 120
Qy 915 AAAAAAAAA-AGTGAATGCTTTTGGAGCAAGTACTCTTCTTCTTCTTCTT 973
Db 121 AAAAAAAAAAGTGAATGCTTTTGGAGCAAGTACTCTTCTTCTTCTTCTT 179
Qy 974 TTTTCTTTTCAAAAAATCTCTAGATTGAATCTTGGAAATGGCTTAAGTCTTCTC 1033
Db 180 TTTTCTTTTCAAAAAATCTCTAGATTGAATCTTGGAAATGGCTTAAGTCTTCTC 239
Qy 1034 TTGTGGCAATTTTGAATGAAAAATACATGTCATTAATTAATCTGAACTTTTAA 1093
Db 240 TTGTGGCAATTTTGAATGAAAAATACATGTCATTAATTAATCTGAACTTTTAA 299
Qy 1094 AAAACCATCATGAGGTTCAAAATATCAAAATATTAATTTGTGTGATATAGACATA 1153
Db 300 AAAACCATCATGAGGTTCAAAATATCAAAATATTAATTTGTGTGATATAGACATA 359
Qy 1154 ACTCTTAATTTTCCCTTAATAATGATTTGTTATATATCTTCAATTCCTCTCATTAT 1213
Db 360 ACTCTTAATTTTCCCTTAATAATGATTTGTTATATATCTTCAATTCCTCTCATTAT 419
Qy 1214 GATTGATATATATAGTGGCAATATCTTAGGAATCTAAAGAGAAAGTTGCAATTG 1273
Db 420 GATTGATATATATAGTGGCAATATCTTAGGAATCTAAAGAGAAAGTTGCAATTG 479
Qy 1274 AAGACTACAGACTGCAAAACCAATTTAAGCCAGATTCCTTGACATGTTGCTTTAATAT 1333
Db 480 AAGACTACAGACTGCAAAACCAATTTAAGCCAGATTCCTTGACATGTTGCTTTAATAT 539
Qy 1334 AGTACTTTACATATAGTAATAACATTATATCATATATGTGGAAGAGCAAGCAAGAAAG 1393
Db 540 AGTACTTTACATATAGTAATAACATTATATCATATATGTGGAAGAGCAAGCAAGAAAG 599
Qy 1394 AAGAAAGTATTTTCATTCAAATCTCTCTCATCAATGCGTAATATCATCATTTG 1453
Db 600 AAGAAAGTATTTTCATTCAAATCTCTCTCATCAATGCGTAATATCATCATTTG 659
Qy 1454 TACAGTTAAGAACCAATAGTGCTGCACACATAGTTTGTGAATTAATGAATGAA 1508
Db 660 GACAGTTAAGAACCAATAGTGCTGCACACATAGTTTGTGAATTAATGAATGAA 714

RESULT 5
US-10-450-826-114
; Sequence 114, Application US/10450826
; Publication No. US20040101818A1
; GENERAL INFORMATION:
; APPLICANT: Ji, Darren
; APPLICANT: Axelrod, Douglas W.
; APPLICANT: Cook, Jonathan S.
; APPLICANT: Jaiswal, Neilam
; APPLICANT: Bistefni, Richard
; APPLICANT: Houghton, Adam
; APPLICANT: Metz, Lawrence
; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation
; FILE REFERENCE: 044921-5039-NO
; CURRENT APPLICATION NUMBER: US/10/450,826
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 60/255,882

Db	15501	ATGCTTCCTTAGAAGAAATGTTTCAGTTAAAGTATTAGAGCCGGAGCCGGTGCCTCA	15442
QY	643	CGCCTGTAAATCCAGCACTTTGGAGGCCGAGGCGGGTGGATCAAGA GTTAGAGATCG	702
Db	15441	CCCCCTGTAAATCCAGCAACCTTGGGGTGGCCGAGGGGGGCAATACAGAGGTCAAGAAATCA	15382
QY	703	AGACCAATCCGGGCTTAACAAGGTGAATAACCTGTCTCTTAATAAAATCAAAAAATTAGCCG	762
Db	15381	AGACCTTCCTGGCTTAACAATGGTGAATAACCCGTCCTCACAAAAAATACAAAAAAATTAGCCG	15322
QY	763	GGTCGCGGTGCAGACGCGCTGTAGTCCAGCACTGCTCAGAGAGCTGAAGCAGAGAAATGGTG	822
Db	15321	GGCGTGGTGGCGGGCGGCTGTAGTCCAGCTTACTCAGGAAGCTGAGAGCAGGAATAGCGG	15262
QY	823	TGAACCCCGGAGAGGTGGAGCTTTCAGTGAAGCCGAGATTGGCGCACTGCACTCCAGCTGGG	882
Db	15261	TGAACCCCAAGGAGGGGAGCTTTCAGTGAAGCCGAGATCGCACCTGTACTCCAGCTGGG	15202
QY	883	CGACAGTGCAGACTGTCTCTCAAAAAAATTTTTTTTTTTTTTTAATG	932
Db	15201	TGAACAGAGCAAGACTCCGTCTCAACAAAGAAAAAATTTAAGTTTG	15152

RESULT 7

US-10-450-826-3/C
Sequence 3, Application US/10450826
Publication No. US20040101818A1
GENERAL INFORMATION:
APPLICANT: JI, Darren
APPLICANT: Axelrod, Douglas W.
APPLICANT: Cook, Jonathon S.
APPLICANT: Jaiswal, Neelam
APPLICANT: Eistein, Richard
APPLICANT: Houghton, Adam
APPLICANT: Mettz, Lawrence
TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation
FILE REFERENCE: 044921-5039-WO
CURRENT APPLICATION NUMBER: US/10/450,826
CURRENT FILING DATE: 2003-06-18
PRIOR APPLICATION NUMBER: US 60/255,882
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: US 60/285,691
PRIOR FILING DATE: 2001-04-24
NUMBER OF SEQ ID NOS: 149
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 92638
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. AC004596
US-10-450-826-3

Query Match	7.0%;	Score 347.6;	DB 17;	Length 92638;
-------------	-------	--------------	--------	---------------

Matches 618; Conservative 0; Mismatches 299; Indels 21; Gaps 7;

Qy	5	GGCCAAATGCTGGCTCAACCGTATATTTCCACAGACTTTTGGAGGACG-----GAGATC	59
Db	90650	GCCAAAGTCAGATGGCTCAACCTGTAAATCCAGACATTTTGGAGGTCCAGGAGGATGATC	90591
Qy	60	GCTTGAGCTCAGCAATTGGAGACACAGCCTACGTAACTAGTGAACCTCTGTCTGTACAA	119
Db	90590	ACTAGAGGCCAGAGATTGGAGACCAGCCTGGCCACATGTTGAAACCTCGTCTCTACTAA	90533
Qy	120	ATAATAAAAGATTTTCCAGGCATGTGGCGTGCACCCCAAGTCCACAGTATTTTGGAGG	179
Db	90530	AAATATCAAAAATTGGCTGAGCATGTGGCTGTCAACCTGTAAATCCAGCACTTTGGGAGG	90471
Qy	180	CTGAGTGAAGAGCAATGCTTGAAGCCAGAGATTGAAGCAACAAGCTACGCAACATTAATGAG	239
Db	90470	CCGAGGCAAGGCTGATTCACCTGAGGCTCAGGAGTTTCAGAACCAAGGCTGTCCAACTGTGTGAA	90411

QY	240	ACCGTGCTCATTAATAAATAATTAAGCT-----GGTGTCTTGGCCACAGGCTCCAGG	292
Db	90410	ACCCCATCTCTAATTAATAATTAACAAAATATGGCTGGGGCATGGCATGGTCCACCTCTATATCC	90355
QY	293	TAGCTACTCGGAGAGACTGAGGTGGAGGATCAC-TGAGCCCAAGGAGGCTGAGGCTCCAGT	351
Db	90350	CAGCACACTCCGGAGGCTGAGGCAAGAGAAATCGCTTGAACTCCAGAGAGCAGAGGTTCTGT	90297
QY	352	GAAACGTGATCACCCAGCTGGATTCACGCTTGGAGACAGAGGAGACCCCTGTTCCAAA	411
Db	90290	GAGTAGAGATCGCGCCATTCCATTCACGCTTAGGTGACA-AGAGCGAAACTGTGTCAAAA	90232
QY	412	AAAAAAAAAAAAAAAAATGCAGAAAAAGCATCATTAACCTTGAGCTGGACATACTTT	471
Db	90231	AAAAAACAAAAACAACAACAACAACA-ACAAAAACAACAACAACAACAACAAAAACCA	90173
QY	472	TATGTGATGAATAATCAACATCTTTTGGAGAAATTAGCATTTCTGATTAATGTATTAT	531
Db	90172	AAATATTATCCGGGCGCAGTGGCTCAGCCTGTATCTTAGACACTGTGGAGGCGGAGGT	90113
QY	532	AATATATATTATTAATTAATTCMAATGGAATTAATATTTCTGAGAAACTAGCTTCTCACT	591
Db	90112	GGGTGATTTACTTGAGGTGACGAGATTTCAGAGCGCTGGCCACAAGGTTGAATCCCAT	90053
QY	592	CTCAGTTTGATGCAAAAACCTTTA--ATGGCTTTTGGCCGGTGGCTCAACGCTGT	649
Db	90052	CTTCACTATAAAATAATAAATTAAGCTGATGTGGGCGCGGGTGGCTCAACGCTGT	89933
QY	650	AATCCACGCACTTTGGGAGGCCAGCGGGTGTATACAAAGTTTAGAATATGAAACCT	709
Db	89992	GGTCCACGCACTTTGGAGGCGCAAGGTAGCGTGAATCAGAGGTCAAGAAATGAGACCAT	89933
QY	710	CTTGCTTAACACGGTGAACCTTGTCTTCACTATAA-----ATATCAAAAAAATTAGCCGGGT	765
Db	89932	CTTGCTTAATATGTGTGAACATGTCTCTTAAAAAATAATAATAATAATAATAATTAAGTGGGT	89873
QY	766	GCGGTGCCAAGCGCTGTAGTCTCCAGCTGCTCAGAGGCTGAAGGCAAGAAATGTTGTGA	825
Db	89872	GTTGTGATGTGGCACTGTATACCCCACTACTTGGGAGGCTGAAGGCAAGAAATCGCTTGA	89813
QY	826	ACCCGGAGGCGGAGCTTGCAGATGAGACCGAGATTGGCCACTGTGCACTCCAGCTTGGCGCA	885
Db	89812	ACCCGGAGGCAAGGTTGCAATGAGCCAAAGATCAACACCATGTGCACTCCGCGCTGGGCAA	89753
QY	886	CAGTGCAGACCTGTCTCAAAAAAAAAAAAAAAAAAAAAA	923
Db	89752	CAGAGTAGACTCAAGTCTCAAAAAAAAAATAATATAATAA	89715

RESULT 8

```

US-10-322-281-296/c
/ Sequence 296, Application US/10322281
/ Publication No. US20040126762A1
/ GENERAL INFORMATION:
/ APPLICANT: David W. Morris
/ APPLICANT: Marc S. Malandro
/ TITLE OF INVENTION: Novel Compositions and Methods in Cancer
/ FILE REFERENCE: 52945201000
/ CURRENT APPLICATION NUMBER: US/10/322,281
/ CURRENT FILING DATE: 2002-12-17
/ NUMBER OF SEQ ID NOS: 866
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 226
/ LENGTH: 59748
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)..(59748)
/ OTHER INFORMATION: n = A,T,C or G
US-10-322-281-296

```

Query Match 6.9%; Score 347.2; DB 17; Length 59748;
Best Local Similarity 68.8%; Pred. No. 7,7e-59;
Matches 561; Conservative 0; Mismatches 228; Indels 27; Gaps 5;

126 AAGAAATTTTCAGGAGTGGTGGCGTGGACCCCGAGGCTATTTGGAGAGCTGAG 185
25609 ATGAAAGTGGCCCGGCGCGGTGCTCAGACCTGTATTCAGAGCATTTGGAGAGCCAAAG 25550
186 TAGAGAGAAATCTTGAAGCCAGAGATTGAAGACAGCTTAGGCAACATAGAGACCTTG 245
25549 TGGGAGATCACTTGAAGTCAAGAGATTCAAGACAGGCTGGCAACATGGTGAATCCCA 25490
246 TGTCTATAAAAT-----AATTAGCTGTGTCTTGGCAAGGCTCGAG--CTAGCTA 298
25489 TCTCTATAAAATACAAAATTAAGCGGTGTGTGGCAATGCTCATAGTCCAGCTA 25430
299 CTGGAAAGCTGAGTGGAGAGATCA-CTGAGCCAGAGAGGCTGAGGCTCAGTGAACAG 357
25429 CTGGAGGCTGAGCGGAGAAATCAGCTGAACTGGAGAGGAGAGGTTGCAATGAGCCC 25370
358 TGATCAAGCAGCTGATTTCCAGCTTGAGAGACAGAGAGAGACCTGTTCACAAAAAAA 417
25369 AGATGGCGCATTTGCACTCCAGCTGGGCAAGAGTGAAGACTCCGTCTCAAAAAACAA 25310
418 AAAAAAAAAATGCAAGAAAGACATCATTAACCTGAGCACTTGAATCTTATGTG 477
25309 AACAAAACATTAATGTAAGAAATTAATTAATCTGATAGCAGCTAGGGTATCTA 25250
478 ATGAATTCACAACTTTTGAAGAAATTAATTAATCTGATTAATTAATTAAT 537
25249 TAGTCACAAAATGTGTGTACTCAGGTAGTACT-----CTCCTAGTACCTGAATTT 25194
538 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 597
25193 ATCACTATACAT-----TATATACAGTACAAATTAATTAATTAATTAATTAAT 25149
598 TGTCAGTCAAACTTTTAATGTTCTTTGGCGGAGTGGCTGAGCTGATATCCAG 657
25148 CTTGAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 25089
658 CACTTTGGAGGCGGAGCGGTGATCAAGAGTTAGAGATGAGAACATCTGAGCTA 717
25088 CACTTTGGAGGCGGAGCGGAGCGGATCAGAGGTGAGAGATGAGAACATCTGAGCTA 25029
718 ACAGCGTGAACCTGCTCTCTACTAATAATTAATTAATTAATTAATTAATTAATTAAT 777
25028 ACAGCGTGAACCTGCTCTCTACTAATAATTAATTAATTAATTAATTAATTAATTAAT 24969
778 GCTGTAGTCCAGCTGCTGAGAGGCTGAGGAGAGAAATGAGTGAACCCGAGAGGCG 837
24968 GCTGTAGTCCAGCTGCTGAGAGGCTGAGGAGAGAAATGAGTGAACCCGAGAGGCG 24909
838 GAGCTTGAAGTGAAGCCAGATTTGGCCAGCTGACCTGAGGCGAGTGGAGAGCT 897
24908 GAGCTTGAAGTGAAGCCAGATTTGGCCAGCTGACCTGAGGCGAGTGGAGAGCT 24849
898 CTGTCTCAAAAAAAAAAAAAAAAAAGTTGAATG 933
24848 CCATCTCAAAAAAAAAAGAGAGAGAGAGAGAGG 24813

RESULT 9
US-10-017-161-2233/C
; Sequence 2233, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKISO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 08435/0152
; CURRENT APPLICATION NUMBER: US/10/017.161

CURRENT FILING DATE: 2002-12-18
PRIOR APPLICATION NUMBER: JP 2001/246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2430
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2233
LENGTH: 59588
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: source
LOCATION: (1) .. (59588)
FEATURE:
NAME/KEY: CDS
LOCATION: (201) .. (239)
FEATURE:
NAME/KEY: CDS
LOCATION: (19765) .. (20013)
FEATURE:
NAME/KEY: CDS
LOCATION: (22071) .. (22222)
FEATURE:
NAME/KEY: CDS
LOCATION: (25670) .. (25814)
FEATURE:
NAME/KEY: CDS
LOCATION: (28077) .. (28268)
FEATURE:
NAME/KEY: CDS
LOCATION: (30114) .. (30265)
FEATURE:
NAME/KEY: CDS
LOCATION: (42934) .. (43362)
FEATURE:
NAME/KEY: CDS
LOCATION: (46385) .. (46558)
FEATURE:
NAME/KEY: CDS
LOCATION: (48143) .. (48221)
FEATURE:
NAME/KEY: CDS
LOCATION: (50255) .. (50359)
FEATURE:
NAME/KEY: CDS
LOCATION: (59242) .. (59388)
US-10-017-161-2233

Query Match 6.9%; Score 345.2; DB 15; Length 59588;
Best Local Similarity 68.9%; Pred. No. 1.9e-58;
Matches 554; Conservative 0; Mismatches 203; Indels 47; Gaps 4;

136 CCAGGATGTTGGCCGTCACCCCACTGACCTATTGGAGGCTGAGGTAGAGAAAT 195
46072 CCAGGATGTTGGCCGTCATGCTCTTAATCCAAAGTTGGAGGTCAAGGTGGAGAAAT 46013
196 GCTTGAAGCAGAGTGAAGCAAGCTAGGCAACATAGTAGACCTGTGTCTA-TAA 254
46012 GCTTGAAGCAGAGTGAAGCAAGCTAGGCAACATAGTAGACCTGTGTCTA-TAA 45953
255 AAAATTAATTAAGCTGTGTCTTGGCAGAGGCTGAG--CTAGCTACTGGAAGACTGAG 312
45952 AAAATTAATTAAGCTGTGTCTTGGCAGAGGCTGAG--CTAGCTACTGGAAGACTGAG 45893
313 GTGGAGAGATCAC-TAAGCCAGAGAGCTGAGGCTGAGTGAACAGTATCAACCCAGCTG 371
45892 GTGGAGAGATCACTTGAAGCCAGAGGTGAAGCTATATGAGCTGTGATCACACACTG 45833
372 GATTCAGCTGGAAGACAGAGGAGACCTGTTCACAAAAAATTAATTAATTAATTAATTAAT 431
45832 TACTTCAAGCTGGGCAACAGAGAGACCTGTGTGAAAAAATTAATTAATTAATTAATTAAT 45773
432 CAAGAAAGACATCATTAATCTGAGCTGGAAGTAACTTTATGATGAAATTCACAAAT 491

Db	45772	AGACAAATGATGTTGGCTGCCTGCATGAAATTCAGACACTTTGAGAGCCAAAGCAAGAAAGATC	45732
Qy	492	CTTTTACGAAAGAAATTAGCATTTCTGATTAATAAGTATTAATTAATTAATTAATTC	551
Db	45712	TCTTGAAGCCCAAGAGTTTCGAGACCAGCTTGAAGCAACATAGCAAGACCACTTTCTTAAA	45653
Qy	552	AAATGGAATTAATAATTCTGAGAAACTAGCTTCTCACTCTCAGTTGTACGTCAAAACT	611
Db	45652	AAAAAAAAAAAAAAAAATTAGCA-----	45628
Qy	612	TTTAATGATCTTTGGCCGGGTGCGGTGCTCAAGCTGTAAATCCAGACCTTTTGGAGGCC	671
Db	45627	-----GGCATGCCAGGCGCGGTGCTATCTCTGTAAATCCAGACCTTTTGGAGGCC	45576
Qy	672	GAGCGGGGTGATTCACAAGGTTTAGAGATCGAGACCATCTGTGCTAAACGGTGAACCT	731
Db	45575	GAGCGGGGTGATTCACAAGGATCGAGAGATTAACACATCTGTGCTAAACGGTGAACCC	45516
Qy	732	CGTCTCTAATTAATAATCAAAAAAAAAATTAGCCGGGTGCGGTGCCAGACGCTGTAGTCCAG	791
Db	45515	TGTTCTTAATTAATAATAATCAAAAAAAAAATTAGCTGGCGGTGTGGCGCGGCTGTAGTCCAG	45456
Qy	792	CTGCTCAAGAGGCGTGAAGCAGAGGAATGATGATACCCGGAGAGCGGAGCTTGCAAGTGA	851
Db	45455	CTTACTCAGAGGCTGAAGCAGAGGAAGGGGTAACTTGGAAGCGGAGCTTGCAAGTGA	45396
Qy	852	CCGAGATTGGCCCATCTGCACTCCAGCCTGGGCCAGATGCGAGACTCTGTCAAAAAA	911
Db	45395	CCGAGATTGGCCCATCTGCACTCCAGCCTGGATGACAGATGAGACCTGTGTCAAAAAA	45336
Qy	912	AAAAAAAAAAAAAGTTGAATGTC	935
Db	45335	TAAAAAATAAAAAATAATTAGCC	45312

```

RESULT 10
US-10-292-798-1879/c
/
Sequence 1879, Application US/10292798
Publication No. US20030235833a1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABEHATANI, HIROYUKI
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
FILE REFERENCE: 084335/166
CURRENT APPLICATION NUMBER: US/10/292,798
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 10/017,161
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: JP 2001-246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2070
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1879
LENGTH: 59588
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
LOCATION: source
FEATURE:
LOCATION: (1)..(59588)
FEATURE:
LOCATION: (1)..(239)
FEATURE:
LOCATION: (201)..(239)
NAME/KEY: CDS
NAME/KEY: CDS
LOCATION: (19765)..(20013)
FEATURE:
NAME/KEY: CDS
LOCATION: (22071)..(22222)
FEATURE:
NAME/KEY: CDS

```

Query	Match	Query Match	Score	DB	Length
136	CCAGGCAATGGGCGTGCACCCCGAGTCCAGGCTATTGGGAGGCTGAGTGAAGGAAT	195	6.9%;	345.2;	59588;
46072	CCAGGCACTATGCTCTACATGCTCTCTATTCACAGATTGGGAGGCTCAAGTGGAGGATT	4601	68.9%;	1.9e-58;	
196	GCTTGAAGCAGAGATTGAAGACAGGCTAGGCAACATAGTGAAGCCCTGTCTA-TAA	254	0;	203;	47; Gaps 4;
46012	GCTTGAAGCAGAGATTGAAGACAGGCTAGGCAACATAGTGAAGCTCAATCTCTA	4595			
255	AAATTAATTAGCTGTGTCTTGGCAGAGCTGAG--CTAGTACTCGGAGACTGAG	312			
45952	AAATTAATTAGCTGTGTCTTGGCAGAGCTGAG--CTAGTACTCGGAGACTGAG	4589			
313	GTCGAGAGATCAC-TGAGCCAGGAGGCTGAGCTGAGTGAACAGTATCACCCAGCTG	371			
45892	GTCGAGAGATCAC-TGAGCCAGGAGGCTGAGCTGAGTGAACAGTATCACCCAGCTG	4583			
372	GATCCAGCCTGGAAGACAGAGGAGACCTGTTCCAAAAAATAAATAAATG	431			
45832	TACTTCAGCCTGGAAGACAGAGGAGACCTGTTCCAAAAAATAAATAAATG	4577			
432	CAGAAAAAGACATCAATTAACCTTGAACCTGAGCAATTAATTTATGTAGTAATTCACAT	491			
45772	AGACAAATATGATGTGCTGCTGCTCAATTCAGACATTTGAGAGGCCAGAGGATC	4571			
492	CTTTTAGGAAGAAATTAGCAATTTCTGATTAATAATTAATTAATTAATTAATTC	551			
45712	TCTTGAAGGCAAGAGTTGAGACCAAGCTTGAAGCAATTAACCAAGACCCATTTCTTAA	4565			
552	AAATGAATTAATTAATTTCTGAGAACTAGCTTCACTCTCAGTTGACGTCAAAACT	611			
45652	AAATGAATTAATTAATTTCTGAGAACTAGCTTCACTCTCAGTTGACGTCAAAACT	4562			
612	TTAATGATCTTTGGCCCGGATGCGGTGCTACGCTGTAAATCCAGCACTTTGGAGGCC	671			
45627	-----GGCATGCGCAGGCGCGGTGCTCATGCTGTAAATCCAGCACTTTGGAGGCC	4557			
672	GAGCGGCTGATTCACAGGTTGAGAGTGAAGACCATCTGCTAAACAGGTGAAACT	731			
45575	GAGCGGCTGATTCACAGGTTGAGAGTGAAGACCATCTGCTAAACAGGTGAAACT	4551			
732	CGTCTCACTAAATAATCAAAAAAATTAGCCGGGCGGCTGAGAGCGCTGTAGTCCAG	791			
45515	TGCTCTCACTAAATAATCAAAAAAATTAGCCGGGCGGCTGAGAGCGCTGTAGTCCAG	4545			
792	CTGCTCAAGAGGCTGAGGACAGAAATGATGTGAACCCGGGAGCGGAGCTTGCAGTGA	851			

```
Db 45455 CTACTCAGGAGGCTGAGGCGAGAGAGGCGTGAACCTGGAGCGGAGCTTCAGAGGAG 45396
Qy 852 CCGAGATTGGCGCACTGCACTCCAGCTGGGCGCAAGTGGCAAGCTGTCTCAAAAAA 911
Db 45395 CCGAGATTGGCGCACTGCACTCCAGCTGGATGACAGAGTGAAGACTGTCTCAAAAAA 45336
Qy 912 AAAAAAAAAAAGTTGATGTC 935
Db 45335 TAAAAATATAAAAAATTATGACC 45312

RESULT 11
US-10-271-416-1
; Sequence 1, Application US/10271416
; Publication No. US20040043021A1
; GENERAL INFORMATION:
; APPLICANT: Keith, Tim
; APPLICANT: Little, Randall D.
; APPLICANT: Van Berdewegh, Paul
; APPLICANT: Dupuis, Josee
; APPLICANT: Del Mastro, Richard G.
; APPLICANT: Allen, Kristina
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO RESPIRATORY DISEASES AND OBESITY
; FILE REFERENCE: 2976-4045
; CURRENT APPLICATION NUMBER: US/10/271,416
; PRIOR FILING DATE: 2002-10-11
; PRIOR FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 304905
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-271-416-1

Query Match 6.8%; Score 341.2; DB 16; Length 304905;
Best Local Similarity 67.2%; Pred. No. 2,6e-57;
Matches 625; Conservative 0; Mismatches 253; Indels 52; Gaps 8;

Qy 10 ATCTGTGGCTCAGCGGTGATTCAGAGCACTTGG-----GAGGAGAGAGATGAGCT 63
Db 24379 ATGCAATGGCTATGCTGTATCTGATCTTGGAGCTTGAAGGAGGAGGATACCT 24438
Qy 64 GAGCTCAGGAATTGAGACAGAGCTTACGTAACATAGTGAACCTGTCTGTACAAATTA 123
Db 24439 GAAGTTGGAGGTTGAGAGCTAGAGCTGAGCAATGTGTAAC-CCATCTCCACTAATA 24497
Qy 124 TAAAGAATTTCAGAGCATGTGGGTGTCACCCCAAGTCCAGCTATTTGGAGGCTGA 183
Db 24498 TACAAAAGT-----ATGGTGGCGCACACTGTATCTCCAGCACTTTGGAGAGCA 24548
Qy 184 GGTAGAGAGATGCTTGAAGCCAGAGTTGAAGACAAAGCTTAGGCAATAGTAGACCC 243
Db 24549 GGGGGTGGATCACTCGAGGTGAGAGTTCAAGACCAAGCTTGCACATGGCAAAACC 24608
Qy 244 TGTGTCTATATAAAT-----AATTAGCTGTGTCTTTGGCAAGGCTGCA--GCTAG 295
Db 24609 CGCTCTACTATAAATAATCAAAAAATTAGCTGGGCAATGTGTGGGCACTGTATCCAG 24668
Qy 296 CTACTCGGAAGATGAGGTGGAGGATCAC-TCAGGCCAGAGGCTGAGGCTGCACTGA 354
Db 24669 CTACTCGGAGGCTGGGCAAGAGATCCCTTGAACCCAGAGGCGAGAGGTTGGATGAG 24728
Qy 355 CAGTGTACCCAGCTGTGATTCAGGCTGGAAGACAGAGGAGAGCCCTGTTCCTCAAAA 414
Db 24729 CCGAATTCACACATGCACTCCAGCTGGGTGACAGAGCGAGATTCCTCTCAAAAAA 24788
Qy 415 AAAAAAAAAAATGCAAGAAAAAGATCATTAACCTTGACCTGGAGCATTACTTTAT 474
Db 24789 AAAAAAAAAA-----ATTAGCTGGGGTGTCTGACGATACCTGTAT 24830
```

```
Qy 475 GTGATGAATTCACATCTTTTAGAGAAATTAGCATTTGTATTAATTAAT 534
Db 24831 CCAGCTACTTGGAGGTGAGAGCATGAGACTCA-----CTTGAACCCAGAGAGC 24881
Qy 535 TATATTATTAATTAATTCAAATGAAATTAAATTTCTGAGAAACTAGCTTCTACTC 594
Db 24882 GAGGTGAAGTACATGATTCGATTCATTCGCTCCGCTTGGGGGACAGATAGACT 24941
Qy 595 AGTTGTCAGTCAAACTTTATATGTTCTTTGGCGGGGTGGGTGCTCAGCCTGTAATCC 654
Db 24942 CTGCTCAAAAAAATTAATAAATAAAGCCGCGGCGAGTGGCTCAGCCTGTATATC 25001
Qy 655 CAGCACTTTGGAGGCGGAGGCGGTGATCAAGAGTTAGAGATGCAATCTTG 714
Db 25002 CAGCACTTTGGAGGCGGAGGCGGAGGATCAAGAGTCAAGAGTCAAGATCAACCTGG 25061
Qy 715 CTACACGGTGAACCTGCTCTCTACTAATAATTAATAAATTAACCGGGGTGGTCCCA 774
Db 25062 CTACACGGTGAACCCGCTCTCTAATAAATAAATAAATTAACCGGGGTGGTGGCG 25121
Qy 775 GAGCCTGTATGCCAGCTGCTCAGAGGCTGAGGAGAGAGATGTGTGAACCCGGAG 834
Db 25122 GCGCCTGTATGCCAGCTGCTCAGAGGCTGAGGAGAGATGTGTGAACCCGGAG 25181
Qy 835 GCGAGCTTGCAGTGAAGCCGAGATTGCGCACTTCAGCCTGGGCGACAGTGGAG 894
Db 25182 GCGAGCTTGCAGTGAACCCGAGATGCGCCACTGCATCCAGCCTGGGCGAGAGAGC 25241
Qy 895 ACTGTGTCAAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 924
Db 25242 TCGTCTCAAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 25271

RESULT 12
US-10-087-192-112/C
; Sequence 112, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 52945200012
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 112
; LENGTH: 45864
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(45864)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-112

Query Match 6.8%; Score 340.8; DB 13; Length 45864;
Best Local Similarity 66.2%; Pred. No. 1.3e-57;
Matches 556; Conservative 0; Mismatches 272; Indels 12; Gaps 4;

Qy 92 TAACATAGGAACCTGTCTGTCAATATATAAATAATTTTCAGAGCATGAGGCT 151
Db 18139 TATCACTTAATAAACAGGATATCGTATTTAATAAATTAACACAGCCGCGGTGCTC 18080
Qy 152 GCACCCAGTGCAGCTATTGTGGAGGCTGAGGTGAGAGAAATGCTTGAAGCCAGAGT 211
Db 18079 ACGCCTGTATCTTACACTTTGGAGGCTGAAGTGGCGGATGCTTGAAGTCAAGAGT 18020
```

```

QY 212 TGAAGACAGCCTAGGCAACATAGTGAACCTGTGTCT-----ATTAATAATTAGC 266
DB 18019 TCGAAGACCGCTGGGCAACGTGTGTAAACCGCATCTTACTGAATAACAAAAATTAAGT 17960
QY 267 TGGTGTCTTTGGCAGCAGCCTGAG--CTAGCTACTCGAAGACTAGGTGGAGAGATCA 324
DB 17959 CAGGGGTGGTGGCAGAGTGTCTTATGTCACAGCCTGGGAGGTTGAGGCAACAGATTA 17900
QY 335 CTGAGCCCAAGAGGCTGAGGCTGACAGTGAACAGTATCAACGAGCTGAATTCAGCCTGG 384
DB 17899 CTGTAATATGGAGGCGGAGAGCTGACAGTGAAGTATCAACGACCTGCATCCAGCTGG 17840
QY 385 AAGACAGAGGAGACCCCTGTTTCAAAAAAATTAATAATAATAATAATAATAATAATAATA 444
DB 17839 GCGACACAGCAGAGTGTCTTCAACAAACACACACACACACTTACATTCATTA 17780
QY 445 CATTAACCTGACCTGGGACATTAACCTTTATGTATGAATTAATCAATCTTTAGAGAA 504
DB 17719 TAAAGAAATCAATTTCAAAACCTGCTGAGCTGACACACTAGATATCTCTATGTC 17720
QY 505 ATTACATTTCTGATTAATAATGATTAATAATAATAATAATAATAATAATAATAATAATA 563
DB 17719 ATTCTGTATCAATGCGCAGTTCTCTCTTTCTTCAAAAGAGTGCACCAAGATTTAG 17660
QY 564 ATATTCTGAGAACTAGCTTCTCACTCTCTGAGTGTCTCAAACTTTAATGCTTTT 623
DB 17659 GGAGCCAGTAAGACATCTTTTTCACATTCACATTTGAGCTTAATAATAATAAGACA 17600
QY 624 GGCAGGAGTGGTGGTGTACGCGCTGTATCCAGCAGCTTTGGAGGCGAGGCGGTGGA 683
DB 17599 GCGCCGGGACGGGTGGTCTCA---TGTATTCAGCAGCTTTGGAGGCGAGGCGAGGAGA 17544
QY 684 TCACAAAGTTAGAGATGAGACCAATCTGCTTAACACAGGTGAACCTGCTCTTAATA 743
DB 17543 TCACAAAGTTAGAGATGAGACCAATCTGCTTAACACAGGTGAACCTGCTCTTAATA 17484
QY 744 AATTCAAAAAATTTAGCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 803
DB 17483 AATTCAAAAAATTTAGCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 17424
QY 804 CTGAGGAGGAGAGATGATGTGAACCGGAGGCGAGGCTTGGAGTGAAGCGAGATTGGGC 863
DB 17423 CTGAGGAGGAGAGATGATGTGAACCGGAGGCGAGGCTTGGAGTGAAGCGAGATTGGGC 17364
QY 864 CACTGCATCTCAGCCTGGGCGAGAGTGCAGAGCTGTCTCAAAAAAATTAATAATAATA 923
DB 17363 CACTGCATCTCAGCCTGGGCGAGAGTGCAGAGAGTCCGTCTCAAAAAAATTAATAATA 17304

RESULT 13
US-10-087-192-286
; Sequence 286, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 286
; LENGTH: 87331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

NAME/KEY: misc_feature
; LOCATION: (1)...(87331)
; OTHER INFORMATION: n = A,T,C or G
; US-10-087-192-286

Query Match: 6.7%; Score 337.2; DB 13; Length 87331;
Best Local Similarity 66.0%; Pred. No. 9.4e-57;
Matches 622; Conservative 0; Mismatches 288; Indels 33; Gaps 8;

QY 11 TGTGTGGCTCAAGCGGTGTATTCAGCACTTTGGAGGCT-----AGAGAGATGCTTG 64
DB 5304 TGCAGTGGCTCAATGCTGTATTCCTCCGACCTTTGGAGGCGCAAGGTGGGATCACTTG 5363
QY 65 AGCTCAGGAATTGGAGACAGCCTTACGTAACTAGTAA--CCTCTGTCTTACAAATA 122
DB 5364 AGCTCAGGAATTGGAGACAGCCTTACGTAACTAGTAACTAGTAACTAGTAACTAGTAA 5423
QY 123 ATAAAGAAATTTTCAGGAGTGGGCGGCAACCCCGAGTGCAGCTATTTGGAGGCTG 182
DB 5424 CAATAATTTGAAGCGGCGGTGTGTATTCAGCTGTATTCAGCACTTTGGAGGCGCA 5483
QY 183 AGTGAAGAGAAATGCTTGAAGCAGAGAGTTGAAGCAAGCCTAGGCAACATAGTGAACC 242
DB 5484 AGATGGGCTGA--TCACAAAGTCAAGAGATGAGACCAATCTGCTTAACACAGGTAAATC 5541
QY 243 CTGTGTCTTAAATAAT-----AATTAGCTGTGTCTTGGCAGAGCC 286
DB 5542 CCGTCTCTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 5601
QY 287 TGCAG--CTAGCTACTCGGAAGACTGAGGTGGAGAGATCAC--TGAAGCCAGAGGCTGAG 343
DB 5602 TGTATGCCAGCTTACTCAGAGAGCTGAGAGAGAGATGAGAGAGATGAGAGAGAGAGAG 5661
QY 344 GCTGAGTGAACAGTATCACCCAGCTGATTTCAAGCTGTAAGACAGAGGAGAGACCTG 403
DB 5662 GTTGACAGTGAAGCAGAGATGAGAGCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 5721
QY 404 TTTCAAAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 461
DB 5722 CTTCAAAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 5781
QY 462 ACATTAATCTTTATGTATGAATAATCAATCTTTTGAAGAAATTAAGCATTTCTGATTA 521
DB 5782 CAGGAGGCTGAGAGAGAGATGAGCTTTGGGCGCGGAGGCGAGGTTTACAGTGAACAG 5841
QY 522 AATGATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 581
DB 5842 ATGAGAGCACTGAATCCAGCCTGGGCGACAGAGATTTTGTCTCAAAATTAATAAT 5901
QY 582 TTCTCACTCTCTAGT--TGTCACTCAAACTTTAATGCTCTTTGCGCGGCTCGGTGCT 640
DB 5902 AATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 5961
QY 641 CAGGCTGTAAATCCAGACCTTTGGAGGCGGAGGCTGATCAAAAGTTAGAGAT 700
DB 5962 CAGGCTGTAAATCCAGACCTTTGGAGGCGGAGGCTGATCAAAAGTTAGAGAT 7021
QY 701 CGAGACATCTGCTGAACAGCTGAACCTGCTCTTAATAATAATAATAATAATAATAATA 760
DB 6022 CGAGACATCTGCTGAACAGCTGAACCTGCTCTTAATAATAATAATAATAATAATAATA 6081
QY 761 CGGAGTGGTGGCAGACCGCTGTATGTCAGCTGCTCAGAGAGCTGAGGAGAGATGG 820
DB 6082 TGGGGGTGGTGGTGGGCGCTGTATATCTCAGCTATCTCGGAGAGATGAAGCAGAGATTA 6141
QY 821 TGTGAACCGGAGGCGGAGGCTTGGAGTGAAGCGAGATTTGGGCACTGACCTCAGCCTG 880
DB 6142 CGTGAACCGGAGGCGGAGGCTTGGAGTGAAGCGAGATTTGGGCACTGACCTCAGCCTG 6201
QY 881 GCGCAGAGTGCAGAGCTGTCTCAAAAAAATTAATAATAATAATAATAATAATAATAATA 922
DB 6202 GCGCAGAGTGCAGAGCTGTCTCAAAAAAATTAATAATAATAATAATAATAATAATAATA 6243

```

```
RESULT 14
US-10-433-580-2/c
/ Sequence 2, Application US/10433580
/ Publication No. US20040091497A1
/ GENERAL INFORMATION:
/ APPLICANT: GENSET
/ TITLE OF INVENTION: SCHIZOPHRENIA RELATED VOLTAGE-GATED ION CHANNEL GENE AND PROTEIN
/ FILE REFERENCE: 93.WO1
/ CURRENT APPLICATION NUMBER: US/10/433,580
/ CURRENT FILING DATE: 2003-06-04
/ PRIOR APPLICATION NUMBER: US 60/251,317
/ PRIOR FILING DATE: 2000-12-05
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: Patent.pm
/ SEQ ID NO 2
/ LENGTH: 237961
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 43726..43868
/ OTHER INFORMATION: exon 8
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 43998..44102
/ OTHER INFORMATION: exon 9
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 52093..52179
/ OTHER INFORMATION: exon 10
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 77568..77699
/ OTHER INFORMATION: exon 11
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 98226..98393
/ OTHER INFORMATION: exon 12
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 106567..106758
/ OTHER INFORMATION: exon 13
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 144109..144246
/ OTHER INFORMATION: exon 14
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 159794..159868
/ OTHER INFORMATION: exon 15
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 191292..191428
/ OTHER INFORMATION: exon 16
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 192967..193108
/ OTHER INFORMATION: exon 17
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 211540..211613
/ OTHER INFORMATION: exon 18
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 225006..225107
/ OTHER INFORMATION: exon 19
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 225544..225613
/ OTHER INFORMATION: exon 20
/ FEATURE:
/ NAME/KEY: exon
```

```
/ LOCATION: 228450..228541
/ OTHER INFORMATION: exon 21
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 228630..228752
/ OTHER INFORMATION: exon 22
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 231289..231345
/ OTHER INFORMATION: exon 23
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 231589..231709
/ OTHER INFORMATION: exon 24
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 231813..231944
/ OTHER INFORMATION: exon 25
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 232900..233067
/ OTHER INFORMATION: exon 26
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 235355..235459
/ OTHER INFORMATION: exon 27
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 51090
/ OTHER INFORMATION: 99-79335-60 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 61293
/ OTHER INFORMATION: 99-79336-369 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 80602
/ OTHER INFORMATION: 99-79338-332 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 100485
/ OTHER INFORMATION: 99-79314-201 : polymorphic base G or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 100509
/ OTHER INFORMATION: 99-79314-225 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 106725
/ OTHER INFORMATION: 99-79316-158 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 166087
/ OTHER INFORMATION: 99-79322-224 : polymorphic base G or T
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 166336
/ OTHER INFORMATION: 99-79322-473 : polymorphic base A or G
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 235894
/ OTHER INFORMATION: 99-79306-182 : polymorphic base C or T
/ FEATURE:
/ NAME/KEY: primer_bind
/ LOCATION: 51031..51051
/ OTHER INFORMATION: 99-79335.pu
/ FEATURE:
/ NAME/KEY: primer_bind
/ LOCATION: 51539..51559
/ OTHER INFORMATION: 99-79335.tp complement
/ FEATURE:
/ NAME/KEY: primer_bind
/ LOCATION: 60925..60945
```

```
OTHER INFORMATION: 99-79336.pu
FEATURE:
NAME/KEY: primer_bind
LOCATION: 61354..61374
OTHER INFORMATION: 99-79336.tp complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 80271..80290
OTHER INFORMATION: 99-79338.pu
FEATURE:
NAME/KEY: primer_bind
LOCATION: 80700..80720
OTHER INFORMATION: 99-79338.tp complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 91037..91056
OTHER INFORMATION: 99-79339.pu
FEATURE:
NAME/KEY: primer_bind
LOCATION: 91466..91486
OTHER INFORMATION: 99-79339.tp complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 100285..100305
OTHER INFORMATION: 99-79314.pu
FEATURE:
NAME/KEY: primer_bind
LOCATION: 100764..100784
OTHER INFORMATION: 99-79314.tp complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 106568..106585
OTHER INFORMATION: 99-79316.pu
FEATURE:
NAME/KEY: primer_bind
LOCATION: 107000..107020
OTHER INFORMATION: 99-79316.tp complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 165864..165884
OTHER INFORMATION: 99-79332.pu
FEATURE:
NAME/KEY: primer_bind
LOCATION: 166381..166401
OTHER INFORMATION: 99-79322.tp complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 235713..235732
OTHER INFORMATION: 99-79306.pu
FEATURE:
NAME/KEY: primer_bind
LOCATION: 236190..236210
OTHER INFORMATION: 99-79306.tp complement
NAME/KEY: primer_bind
LOCATION: 51071..51089
OTHER INFORMATION: 99-79335-60.mls

Query Match 6.7%; Score 336.8; DB 16; Length 237961;
Best Local Similarity 66.3%; Pred. No. 1.8e-56;
Matches 556; Conservative 0; Mismatches 257; Indels 25; Gaps 4;

QY 117 CAATATATAAAAGATTTTCCAGGCGATGTGGCGTGCACCCCGAGTCCAGCTAATTGGG 176
DB 82061 CAAAACACAGTGCAGCGCCAGAGTGTGCTCAGCTGTAATTCAGCACTCTGGG 82002

QY 177 AGGCTAGGTAGAGGAATGCTTGAAGCCAGAGTGAAGCAAGCTAGGCAACAATGT 236
DB 82001 AGGTCAAGATGATGATGATGCTGAGCCGAGGTTTGAGACCGCTGCTTAACATGCG 81942

QY 237 GAGACCTGTGTCTATATAAAATTAATAGTGTGTCTTGGACAGGCGCTGACGTAGC 296
DB 81941 GAAACCCCATCTCTACTAGAAAATACAA---AATTAGCTGGGTGCGCTGTATCCAGC 81885
```

```
QY 297 TACTCGAAAGACTAGAGTGGAGATCACT-GAGCCAGAGGCTGAGGCTGAGTGAAC 355
DB 81884 TTCTTGGATGTGCTAAGGCAAGGAACCACTGGAATCTGGAGGCAAGAGTTGGAGAGC 81825

QY 356 AGTATCACCACGCTGATTCAGGCTGGAAGCAGAGGGAAGCCCTGTTTCCAAAAA 415
DB 81824 CGAGATGTGCACATGCACTTGAAGCTGAGCAACAGAGTGGAGCTGTCTCAAAAAA 81765

QY 416 AAAAAAAAAAAAAATGCAAGAAAGACATCA-----TAACTTGACCTG 459
DB 81764 AAAAAAAAAAAAAAGAGACAAAAAACAACAACAAAAAACCCTAGAGTGAACAA 81705

QY 460 GGAATTAATCTTTATGTATGAATTCACAATCTTTAGAGAAATTAAGATTGTGAT 519
DB 81704 ATTCAGAAATGCTTAAGTATGTAAGTACAGTGAAGGAGTCTATCTGTTGTTAAGTC 81645

QY 520 AAAATGATTAATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 574
DB 81644 AACAAATACACCTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 81585

QY 575 AACTAGCTTCTGACTCTCTAGTGTGCACTTAATTAATTAATTAATTAATTAATTAAT 634
DB 81584 TTTTAGCTGTTAATGATTTTGGCGGTTGTTAAGAAATGAATGACATGTGGCGGCGG 81525

QY 635 GTGGCTACGCTGTATATCCAGCACTTTGGAGGCCGAGGCGGAGTACACAGGTTA 694
DB 81524 GTGGCTACGCTGTATATCCAGCACTTTGGAGGCCGAGGCGGAGTACACAGGTTA 81465

QY 695 GGAGATCGAGACATCTGCTTAACACGGTGAACCTGCTCTATAAATAATACAAAA 754
DB 81464 GGAGATCGAGACATCTGCTTAACACGGTGAACCTGCTCTATAAATAATACAAAA 81405

QY 755 ATTAGCGGCGTGGTGCAGACGCTGTAGTCCAGCTGTACAGAGGCTGAGGAGCA 814
DB 81404 ATTAGCGGCGTGAATGTGGGCGGCTGTAGTCCAGCTGTACAGAGGCTGAGGAGCA 81345

QY 815 GAATGCTGGAACCGGAGGCGGAGCTTGCAGTGAAGCCGAGATTGCCCACTGCACCTCC 874
DB 81344 GAATGCTGGAACCGGAGGCGGAGCTTGCAGTGAAGCCGAGATTGCCCACTGCACCTCC 81285

QY 875 AGCTGGGGGACAGTGCAGACTGTCTCAAAAAAAGTTGAATG 932
DB 81284 AGCTGGGGGACAGTGCAGACTGTCTCAAAAAAAGTTGAATG 81227

RESULT 15
US-10-087-192-1546
Sequence 1546, Application US/10087192
Publication No. US20020182586A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Engelhard, Eric K.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
FILE REFERENCE: 529452000122
CURRENT APPLICATION NUMBER: US/10/087,192
PRIOR FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 2059
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1546
LENGTH: 26444
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(26444)
OTHER INFORMATION: n = A, T, C or G
US-10-087-192-1546
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2004, 17:41:01, Search time 318.254 Seconds
(without alignments)
13849.312 Million cell updates/sec

Title: US-10-786-065-3_COPY_20000_26200

Perfect score: 6201
Sequence: 1 caccatgggctcaggaact.....aagatgaagaatgctg 6201.

Scoring table:
IDENTITY_NUC
Gapop 10_0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6201	100.0	53332	4	US-09-801-861-3
2	6201	100.0	53332	4	US-10-224-562-3
3	402.6	6.5	1930	1	US-08-544-900-1
4	402.6	6.5	1930	5	PCT-US95-07874-1
5	327.6	5.3	128779	4	US-09-497-855A-38
6	306	4.9	22481	3	US-08-167-841A-43
7	306	4.9	22481	5	PCT-US95-07201-43
8	306	4.9	22484	3	US-09-875-223-2
9	306	4.9	22484	4	US-09-875-114-2
10	305	4.9	87350	3	US-08-781-891-79
11	305	4.9	87350	4	US-09-618-166-79
12	304.4	4.9	14581	4	US-08-520-373D-4
13	296.8	4.8	33769	4	US-09-544-398B-8
14	296.8	4.7	33769	4	US-09-543-771-8
15	291.6	4.7	87543	4	US-09-791-211-3
16	288.6	4.7	3609	4	US-09-705-299-11
17	288.6	4.7	63588	4	US-09-873-404-3
18	288.6	4.7	63588	4	US-10-243-735-3
19	283.8	4.6	53526	3	US-08-658-136-2
20	283.8	4.6	53527	3	US-08-658-136-2
21	282.2	4.6	137000	4	US-10-172-911-11
22	282	4.5	63000	4	US-09-780-172-18
23	280.4	4.5	15297	3	US-09-817-180-3
24	280.4	4.5	15297	4	US-10-003-295-3
25	280.4	4.5	114793	4	US-10-148-806-3
26	279.6	4.5	1000	3	US-09-018-584A-32
27	279.6	4.5	1000	4	US-09-784-423-32

28	279.2	4.5	62804	3	US-09-800-960-3	Sequence 3, Appl
29	279.2	4.5	62804	4	US-10-096-960-3	Sequence 3, Appl
30	277.2	4.5	55827	4	US-09-813-133A-3	Sequence 3, Appl
31	275.4	4.4	66804	4	US-09-740-041-3	Sequence 1, Appl
32	275.4	4.4	162450	3	US-09-345-882-1	Sequence 1, Appl
33	275	4.4	39982	4	US-09-820-924-3	Sequence 1, Appl
34	275	4.4	39982	4	US-10-369-628-3	Sequence 3, Appl
35	274.4	4.4	14796	3	US-08-975-080-35	Sequence 35, Appl
36	274.4	4.4	14796	3	US-09-630-706-10	Sequence 10, Appl
37	274.4	4.4	14796	3	US-09-496-694B-3	Sequence 3, Appl
38	273.4	4.4	90541	4	US-09-759-359A-3	Sequence 3, Appl
39	273.4	4.4	90541	4	US-10-207-973-3	Sequence 3, Appl
40	273.2	4.4	26928	4	US-09-544-398B-6	Sequence 6, Appl
41	273.2	4.4	26928	4	US-09-543-771-6	Sequence 6, Appl
42	271.2	4.4	62804	3	US-09-800-960-3	Sequence 3, Appl
43	271.2	4.4	62804	4	US-10-096-960-3	Sequence 3, Appl
44	271	4.4	114793	4	US-10-148-806-3	Sequence 3, Appl
45	270.8	4.4	162450	3	US-09-345-882-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1									
US-09-801-861-3									
Sequence 3, Application US/09801861									
Patent No. 6492154									
GENERAL INFORMATION:									
APPLICANT: YAN, Chunhua et al.									
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC									
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES									
FILE REFERENCE: CLO01098									
CURRENT APPLICATION NUMBER: US/09/801,861									
CURRENT FILING DATE: 2001-03-09									
NUMBER OF SEQ ID NOS: 10									
SOFTWARE: FastSeq for Windows Version 4.0									
SEQ ID NO 3									
LENGTH: 53332									
TYPE: DNA									
ORGANISM: Human									
US-09-801-861-3									
Query Match									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 6201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	CAACATGGGCTCAGGCAATTTATGAAGCCAAATATACAGCCTTAATAATGATGTGAC	60						
DB	20000	CAACATGGGCTCAGGCAATTTATGAAGCCAAATATACAGCCTTAATAATGATGTGAC	20059						
QY	61	CTAATATCCCAAGAGAACTCCCTTTGTAGATTTGTACAAATAATTAATATGATGAG	120						
DB	20060	CTAATATCCCAAGAGAACTCCCTTTGTAGATTTGTACAAATAATTAATATGATGAG	20119						
QY	121	TTAATAGTTCTATGATGATGTGAACCCAGAGCCATATCAGCCTAGCAAAATGGCACA	180						
DB	20120	TTAATAGTTCTATGATGATGTGAACCCAGAGCCATATCAGCCTAGCAAAATGGCACA	20179						
QY	181	ATTCAATATCATCAAAAGTTATCTTCAAGAGCTTCAGGCGCTTAATGATGTAAAGAA	240						
DB	20180	ATTCAATATCATCAAAAGTTATCTTCAAGAGCTTCAGGCGCTTAATGATGTAAAGAA	20239						
QY	241	ATGTGAAGAGCCCTCAGCCATCTGAAGACAGTGTATACAGATGATCAAAAGAAAA	300						
DB	20240	ATGTGAAGAGCCCTCAGCCATCTGAAGACAGTGTATACAGATGATCAAAAGAAAA	20299						
QY	301	CCACAGGCGCTTCCCTTCCCAATCTGATGATGAAGAGCTTCAATTTCCATAGTAGT	360						
DB	20300	CCACAGGCGCTTCCCTTCCCAATCTGATGATGAAGAGCTTCAATTTCCATAGTAGT	20359						
QY	361	AAATTTTCAGATACAGCTGTAGAGCTCAAGTCTGGAAGAAAGATCCCATCAAG	420						

```
Db 20360 AATTTCCTAGATACAGCTTGTAGAGCTCAAGTACTGGAAGAAAGCTCCATTGCAAG 20419
QY 421 GAAATTTATCTTAAGATCTGTAATGATATCTAATTTTGTACATTGGAAATATATAGT 480
Db 20420 GAAATTTATCTTAAGATCTGTAATGATATCTAATTTTGTACATTGGAAATATATAGT 20479
QY 481 TGTTCACCTGGGCGGCTGCAAGCTGTAAATCCAGAGCCCTTTGGAGGCGAGAGTGG 540
Db 20480 TGTTCACCTGGGCGGCTGCAAGCTGTAAATCCAGAGCCCTTTGGAGGCGAGAGTGG 20539
QY 541 CAGATCATGAGTCAAGAGATTGAGACCAAGCTAGCCACATGATGAAAACCCGCTCTTA 600
Db 20540 CAGATCATGAGTCAAGAGATTGAGACCAAGCTAGCCACATGATGAAAACCCGCTCTTA 20599
QY 601 CTAAAGATACAAAAATTAGCCAGTGTGGCGCACCTGTAAACCCAGCTGCTCGA 660
Db 20600 CTAAAGATACAAAAATTAGCCAGTGTGGCGCACCTGTAAACCCAGCTGCTCGA 20659
QY 661 GAGAGTGAAGCAGAGAAATTGCTGAACCCAGAGGAGAGGTGACGCGAGCAAGATCA 720
Db 20660 GAGAGTGAAGCAGAGAAATTGCTGAACCCAGAGGAGAGGTGACGCGAGCAAGATCA 20719
QY 721 CACCAATGCACTGTAGCTGTAGATGACAGGCAAGACTCCAATCAAAAAA 780
Db 20720 CACCAATGCACTGTAGCTGTAGATGACAGGCAAGACTCCAATCAAAAAA 20779
QY 781 AAGAAATATGTAAAGTTGTGCTATTAACAATTAATAGGCACTGAGAACCAAGTCTTAAG 840
Db 20780 AAGAAATATGTAAAGTTGTGCTATTAACAATTAATAGGCACTGAGAACCAAGTCTTAAG 20839
QY 841 CCTATGACCACTGTAACTAGGAATACGTGGGAAACATATTAAGGAACTTAACCCAGT 900
Db 20840 CCTATGACCACTGTAACTAGGAATACGTGGGAAACATATTAAGGAACTTAACCCAGT 20899
QY 901 CCTGAAGTAAAGCTTTTGGAAAGAAATGTTTGAAGA CAAAGGCTTAAGAGAGTAAAAA 960
Db 20900 CCTGAAGTAAAGCTTTTGGAAAGAAATGTTTGAAGA CAAAGGCTTAAGAGAGTAAAAA 20959
QY 961 AAAAAATTAATATCCAGTTTACGTGTGTGAGAAATGGGATAGGAGCTTAATGAGAAAT 1020
Db 20960 AAAAAATTAATATCCAGTTTACGTGTGTGAGAAATGGGATAGGAGCTTAATGAGAAAT 21019
QY 1021 CAAATGAGATGTTTATGATGATGTTAAGGACCTGGTAAAGGTTGAAGACATTAATTA 1080
Db 21020 CAAATGAGATGTTTATGATGATGTTAAGGACCTGGTAAAGGTTGAAGACATTAATTA 21079
QY 1081 TCTGCAACAATCCGCGGACCTTTTATGTGTATGCTTGGCAATTTAATAGAGAGCA 1140
Db 21080 TCTGCAACAATCCGCGGACCTTTTATGTGTATGCTTGGCAATTTAATAGAGAGCA 21139
QY 1141 GAGAAATGATGACGTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 21140 GAGAAATGATGACGTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 21199
QY 1201 AGACAGTGAAGCAGAGAGATTGAAGATTAATCAAGACAGAAATTAATGATGATGATGAT 1260
Db 21200 AGACAGTGAAGCAGAGAGATTGAAGATTAATCAAGACAGAAATTAATGATGATGATGAT 21259
QY 1261 GGCATCTAGCTGAGCTTAATCTGAGGGAAGAAATGAAGATTAAGCAAGCTTGTGATAGT 1320
Db 21260 GGCATCTAGCTGAGCTTAATCTGAGGGAAGAAATGAAGATTAAGCAAGCTTGTGATAGT 21319
QY 1321 TATGAAGAGAGTGAAGGCTTCAAGACCTACAGGTTGATTAATTAAGAAATGATTTG 1380
Db 21320 TATGAAGAGAGTGAAGGCTTCAAGACCTACAGGTTGATTAATTAAGAAATGATTTG 21379
QY 1381 GAGAAAGATTAATCTGAGAGAGTGAAGATTTTCAAGGCTTGAAGATCTACATACAG 1440
Db 21380 GAGAAAGATTAATCTGAGAGAGTGAAGATTTTCAAGGCTTGAAGATCTACATACAG 21439
QY 1441 CACTGTGCTAAATGCTTCAAAAGCATGATCCCTGCTCAAGGAGCTTAACGCAAAAAC 1500
Db 21440 CACTGTGCTAAATGCTTCAAAAGCATGATCCCTGCTCAAGGAGCTTAACGCAAAAAC 21499
```

```
QY 1501 AAGAGATTAAGAAATACACCAATACTATTATAGACACTTGTGTGAATATCAAGAAAG 1560
Db 21500 AAGAGATTAAGAAATACACCAATACTATTATAGACACTTGTGTGAATATCAAGAAAG 21559
QY 1561 AATATAGATCTAGTACTGTAGATGTGCAACGGGATCAAGATATCTTCTAGTTTCAAGAA 1620
Db 21560 AATATAGATCTAGTACTGTAGATGTGCAACGGGATCAAGATATCTTCTAGTTTCAAGAA 21619
QY 1621 GTTTCAGATCGGCGCGGCTGCTACGCTGTAAATCCAGACCTTTGGAGGCGCA 1680
Db 21620 GTTTCAGATCGGCGCGGCTGCTACGCTGTAAATCCAGACCTTTGGAGGCGCA 21679
QY 1681 GCGGCGTGAATACAAAGTCAAGAGATCAAGACATCTCTGTAAACACGCTGAACCCCG 1740
Db 21680 GCGGCGTGAATACAAAGTCAAGAGATCAAGACATCTCTGTAAACACGCTGAACCCCG 21739
QY 1741 TCTCTACAAAAATATATATATATAGCCAGGCGTGTGGCGGCGCTGTAGTCCAGCT 1800
Db 21740 TCTCTACAAAAATATATATATATAGCCAGGCGTGTGGCGGCGCTGTAGTCCAGCT 21799
QY 1801 ACTCAGAGGCTGAGGCAAGAAATGGCGTGAACCCGAGGATGAAGTTTGCCTGAGCCG 1860
Db 21800 ACTCAGAGGCTGAGGCAAGAAATGGCGTGAACCCGAGGATGAAGTTTGCCTGAGCCG 21859
QY 1861 AGATGCGGCTACTGCGCTCCAGCTGGGCGACAGATGAGATCTGCTCTCAAAAAA 1920
Db 21860 AGATGCGGCTACTGCGCTCCAGCTGGGCGACAGATGAGATCTGCTCTCAAAAAA 21919
QY 1921 AAAAAAATTAATAGTTTCAAGATCTTAACACATGATTTCAACAGCTTAGAAATAGAG 1980
Db 21920 AAAAAAATTAATAGTTTCAAGATCTTAACACATGATTTCAACAGCTTAGAAATAGAG 21979
QY 1981 AGCATGTTACAGGAGAGAAAAATGTTTTCAGCAAAAGTACAGATGAGGAAATAGAGAT 2040
Db 21980 AGCATGTTACAGGAGAGAAAAATGTTTTCAGCAAAAGTACAGATGAGGAAATAGAGAT 22039
QY 2041 ATGTTCAAGAGAGAGACCCGAGAGTCAAGTGTGTTGGTGGTTAGAGAAACAGTGT 2100
Db 22040 ATGTTCAAGAGAGAGACCCGAGAGTCAAGTGTGTTGGTGGTTAGAGAAACAGTGT 22099
QY 2101 TTGCAATCTCCAGGTTTCCATTAAGTGTGATGAATCAATATGATGATGATGATGATGAT 2160
Db 22100 TTGCAATCTCCAGGTTTCCATTAAGTGTGATGAATCAATATGATGATGATGATGATGAT 22159
QY 2161 TTTTAAAAATGAATTAATGATGAGAAAGAAATGAATTAATGATGATGATGATGATGAT 2220
Db 22160 TTTTAAAAATGAATTAATGATGAGAAAGAAATGAATTAATGATGATGATGATGATGAT 22219
QY 2221 TTGAAAGAGCAATATATTTTCTGCACTTTTCTGCACTTTTCTGCACTTTTCTGCACTTT 2280
Db 22220 TTGAAAGAGCAATATATTTTCTGCACTTTTCTGCACTTTTCTGCACTTTTCTGCACTTT 22279
QY 2281 TTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
Db 22280 TTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 22339
QY 2341 AATCTTTAAAGCACTGGCTGTGCTTAATCTTATTTTTCAGAGAGAAATCAAGATC 2400
Db 22340 AATCTTTAAAGCACTGGCTGTGCTTAATCTTATTTTTCAGAGAGAAATCAAGATC 22399
QY 2401 TGAAGAGCAAAACATTTTTCGAGAGTTATGAACAGCTTATGCAATGCTTAAGAAAGTA 2460
Db 22400 TGAAGAGCAAAACATTTTTCGAGAGTTATGAACAGCTTATGCAATGCTTAAGAAAGTA 22459
QY 2461 TTCTTAGTTAAATCTTTTCCACTAGTGCATCTAGTGCATCTAGTGCATCTAGTGCATCT 2520
Db 22460 TTCTTAGTTAAATCTTTTCCACTAGTGCATCTAGTGCATCTAGTGCATCTAGTGCATCT 22519
QY 2521 AATACAGAAATATATAGTGAACACAGATACACAGATCTGGGAGAAATATATGGGTAGTG 2580
Db 22520 AATACAGAAATATATAGTGAACACAGATACACAGATCTGGGAGAAATATATGGGTAGTG 22579
```

```
QY 2581 GCTGAGAGCTCAATTTCTAAGAAATGTGACCTTAGGCAAGGTAATGCTCACAACCT 2640
DB 22580 GCTGAGAGCTCAATTTCTAAGAAATGTGACCTTAGGCAAGGTAATGCTCACAACCT 22639
QY 2641 ATAAATCCAGACATTTGGAGGCGCAAGTCAAGAAATCGCTTGAACCAAGATTCAAG 2700
DB 22640 ATAAATCCAGACATTTGGAGGCGCAAGTCAAGAAATCGCTTGAACCAAGATTCAAG 22699
QY 2701 CTAGCATGGGCAACATAGCAAGACTCATCTCAAAAAATTTTAAATCAGCTGACA 2760
DB 22700 CTAGCATGGGCAACATAGCAAGACTCATCTCAAAAAATTTTAAATCAGCTGACA 22759
QY 2761 TGGTGCAATACGCTTGAATCCCACTTACCTGGGAAGCTAGTGGGTGATGCTTGAAC 2820
DB 22760 TGGTGCAATACGCTTGAATCCCACTTACCTGGGAAGCTAGTGGGTGATGCTTGAAC 22819
QY 2821 CAGAGATTTGAGGCTAAGGTGAGCCATGATCACACAATGCACTCCAGCTTGAAGTACAG 2880
DB 22820 CAGAGATTTGAGGCTAAGGTGAGCCATGATCACACAATGCACTCCAGCTTGAAGTACAG 22879
QY 2881 AGGAAGACCTCTGCTCCCTAAAAAAGAAATGTGATTTTATCTTAGACATGACAT 2940
DB 22880 AGGAAGACCTCTGCTCCCTAAAAAAGAAATGTGATTTTATCTTAGACATGACAT 22939
QY 2941 CATTAATCTAATAGTTTGAATGAGAAAAATTAATGATCAGAAAGAAATTTATCACT 3000
DB 22940 CATTAATCTAATAGTTTGAATGAGAAAAATTAATGATCAGAAAGAAATTTATCACT 22999
QY 3001 GTGCTCTGATAGATATATGAAAGAAATAGAGACTAGAGTCAAGGATTCACCTTAAGT 3060
DB 23000 GTGCTCTGATAGATATATGAAAGAAATAGAGACTAGAGTCAAGGATTCACCTTAAGT 23059
QY 3061 TTTGTTGTTTGTGTTTGAAGACAGAGTCTTTTGTGTTACCCAGGCTAGATGCAATGAT 3120
DB 23060 TTTGTTGTTTGTGTTTGAAGACAGAGTCTTTTGTGTTACCCAGGCTAGATGCAATGAT 23119
QY 3121 CAGTCAATGCTCAACCCGACCTCAAACTCCAGCTTCAAAATTTCTTCCAGCTCGGCT 3180
DB 23120 CAGTCAATGCTCAACCCGACCTCAAACTCCAGCTTCAAAATTTCTTCCAGCTCGGCT 23179
QY 3181 CCCAAGCTCTGAAATTAAGGTGAGCCAAAGGTTTATTGATGCTGCTGCTGCTGCT 3240
DB 23180 CCCAAGCTCTGAAATTAAGGTGAGCCAAAGGTTTATTGATGCTGCTGCTGCTGCTGCT 23239
QY 3241 GCCTCTCAAACTTCAAGTGAAGACAGAGTGAACCGGGAACCTGACTCAAGCTGCGGT 3300
DB 23240 GCCTCTCAAACTTCAAGTGAAGACAGAGTGAACCGGGAACCTGACTCAAGCTGCGGT 23299
QY 3301 TAAAGCTGAGCTCTGCAATTTCAATGAGTCAAGCTGATGAGATGCTGCTCAAG 3360
DB 23300 TAAAGCTGAGCTCTGCAATTTCAATGAGTCAAGCTGATGAGATGCTGCTGCTCAAG 23359
QY 3361 GCCAAGCACTGAGCAGCAAGGATCTAGTTAGCAATTAATGATGATGATGATGAT 3420
DB 23360 GCCAAGCACTGAGCAGCAAGGATCTAGTTAGCAATTAATGATGATGATGATGATGAT 23419
QY 3421 AGTGAATTAAGATGAGAGAGATGTGAATCAGTAAACAAAGAGTTCACTCTTG 3480
DB 23420 AGTGAATTAAGATGAGAGAGATGTGAATCAGTAAACAAAGAGTTCACTCTTG 23479
QY 3481 GTAAATGTGACATGAGAGAGGAAAGATGGGCAACATTAATGCTGTTTGTGTTGACT 3540
DB 23480 GTAAATGTGACATGAGAGAGGAAAGATGGGCAACATTAATGCTGTTTGTGTTGACT 23539
QY 3541 GACGAGAGAAATTTGAGCTCTATTTAACAAGAAATAGAGAAAGTTGTTTGAAGAAA 3600
DB 23540 GACGAGAGAAATTTGAGCTCTATTTAACAAGAAATAGAGAAAGTTGTTTGAAGAAA 23599
QY 3601 GAGAGACTCTGTTTCAAGACGTGTGAGGTTCCAGGTGAGACAGAGTCTCCAAAGGAAAT 3660
DB 23600 GAGAGACTCTGTTTCAAGACGTGTGAGGTTCCAGGTGAGACAGAGTCTCCAAAGGAAAT 23659
QY 3661 GAGCAGTAGCAACCTAAAGAAATCTGTGCTCAAGAGGAGCTGTGAGCTGACGTGT 3720
DB 23660 GAGCAGTAGCAACCTAAAGAAATCTGTGCTCAAGAGGAGCTGTGAGCTGACGTGT 3720
DB 23660 GAGCAGTAGCAACCTAAAGAAATCTGTGCTCAAGAGGAGCTGTGAGCTGACGTGT 3720
QY 3721 AGATCTGAGGATCATCAGACATAGAGTTTGAAGACAAAGATGAGCAACAAAGAGC 3780
DB 23720 AGATCTGAGGATCATCAGACATAGAGTTTGAAGACAAAGATGAGCAACAAAGAGC 23779
QY 3781 AAATACAAAGAGAGAGAGTATGATGACATTTTGCCTTTTATGATGAGAAAGGA 3840
DB 23780 AAATACAAAGAGAGAGAGTATGATGACATTTTGCCTTTTATGATGAGAAAGGA 23839
QY 3841 ACAGGAATGAAGGAATGAAGGGAAGCAGCTGTGGAATGTAGAGCATCTGAAAAAAA 3900
DB 23840 ACAGGAATGAAGGAATGAAGGGAAGCAGCTGTGGAATGTAGAGCATCTGAAAAAAA 23899
QY 3901 ATACACACTGTATGAGAGTCAAGGGAAGAAATTTCAAGAAAGAGGATATGTTGACA 3960
DB 23900 ATACACACTGTATGAGAGTCAAGGGAAGAAATTTCAAGAAAGAGGATATGTTGACA 23959
QY 3961 GTATTCAAGCATCAGGAATACAGCTTAAAGTCTATCTTGAATGCAATTTGCTGTG 4020
DB 23960 GTATTCAAGCATCAGGAATACAGCTTAAAGTCTATCTTGAATGCAATTTGCTGTG 24019
QY 4021 ATTTGTGAGGGAACACCTAATTAATTAAGGAATTTATGTTGGGTATATGAGGCAAAAG 4080
DB 24020 ATTTGTGAGGGAACACCTAATTAATTAAGGAATTTATGTTGGGTATATGAGGCAAAAG 24079
QY 4081 GAAGAGTTATCCAAATCAAGCAGGTGAGATGAGGATGAGTTCTCAAGGTGAGGCA 4140
DB 24080 GAAGAGTTATCCAAATCAAGCAGGTGAGATGAGGATGAGTTCTCAAGGTGAGGCA 24139
QY 4141 TCAATGAATGTGGGAAGGGGCAAGAGCATCCATGCCATCCAGGCAAGCCACTTCCA 4200
DB 24140 TCAATGAATGTGGGAAGGGGCAAGAGCATCCATGCCATCCAGGCAAGCCACTTCCA 24199
QY 4201 GAAGCCTCAATAGAGTTAGGATTCAGAGTCTGATACCTTAATCTTCTGAGGTTT 4260
DB 24200 GAAGCCTCAATAGAGTTAGGATTCAGAGTCTGATACCTTAATCTTCTGAGGTTT 24259
QY 4261 TGCATAGGCTTCAATGTTGAGGATGATTAATTAATTAATGAGGCACTGATCAACTT 4320
DB 24260 TGCATAGGCTTCAATGTTGAGGATGATTAATTAATTAATGAGGCACTGATCAACTT 24319
QY 4321 AACCTTCAACCCCTCTCCCTCTCAATATCATGCTTGTGTTTCAAGTACAGTCCCTA 4380
DB 24320 AACCTTCAACCCCTCTCCCTCTCAATATCATGCTTGTGTTTCAAGTACAGTCCCTA 24379
QY 4381 TCCTAAGCTAACCAATGCTGCTGCAAGTCAAGTCAACCTTCAAAAGACATCACTTG 4440
DB 24380 TCCTAAGCTAACCAATGCTGCTGCAAGTCAAGTCAACCTTCAAAAGACATCACTTG 24439
QY 4441 GAGATTTCTAAGATTTTGAAGTGGCTGTCAGAAATTTAGTTGAAGATCAATATATAT 4500
DB 24440 GAGATTTCTAAGATTTTGAAGTGGCTGTCAGAAATTTAGTTGAAGATCAATATATAT 24499
QY 4501 TTCACAATATCAAGTCTGCTATTTATATCAAGGCGCATTAATGTTTAAACAAAG 4560
DB 24500 TTCACAATATCAAGTCTGCTATTTATATCAAGGCGCATTAATGTTTAAACAAAG 24559
QY 4561 AGGTGAATTAATCAGATTTTCTTTTATTAATCACTGATGACAGTGTGATGATG 4620
DB 24560 AGGTGAATTAATCAGATTTTCTTTTATTAATCACTGATGACAGTGTGATGATG 24619
QY 4621 ATTGGATGAGGCAATATCTTTTGTGAAATGTTATATCCCTGACCTTACTTCTC 4680
DB 24620 ATTGGATGAGGCAATATCTTTTGTGAAATGTTATATCCCTGACCTTACTTCTC 24679
QY 4681 CTGTTTCTTCTACCTCTCTCCCTTACTCAACAAGAAATCTTCTCTCTACTCAT 4740
DB 24680 CTGTTTCTTCTACCTCTCTCCCTTACTCAACAAGAAATCTTCTCTCTACTCAT 24739
QY 4741 TCCCTGAATGCTGCTGCTGTTAAGGTTCCAGCTTGAAGTGAAGGCTAATCAGAACGAC 4800
DB 24740 TCCCTGAATGCTGCTGCTGTTAAGGTTCCAGCTTGAAGTGAAGGCTAATCAGAACGAC 4800
```

Db	24740	TTCCCTGAATGCTGGTCTCTGTTAAGGTTCCAGCCTTGAAGAGAGGCTTAATGAAACAC	24792
Qy	4801	AGTGTGACAGATGTGAGATGATGTGGAGAAAGTGTGACAGATATGAGACCAATTACTTA	4860
Db	24800	AGTGTGATACAGATGTGAGATGATGTGGAGAAAGTGTGACAGATATGAGACCAATTACTTA	24859
Qy	4861	GCCGGAACTGACGGGAAAAACAAGATGTGACGCGATATTTTCTGTGATCTGAGTATTAA	4920
Db	24860	GCCGGAACTGACGGGAAAAACAAGATGTGACGCGATATTTTCTGTGATCTGAGTATTAA	24919
Qy	4921	ATGATGATGTGTCCCAATCACTGTGTATGAGAAATCGAAGAAAAATTTATTTTGAGAG	4980
Db	24920	ATGATGATGTGTCCCAATCACTGTGTATGAGAAATCGAAGAAAAATTTATTTTGAGAG	24979
Qy	4981	ATACACATGATTTGTGTTTGAATACCTAGATTTGAGATTTATGAGATGTACAGGGGA	5040
Db	24980	ATACACATGATTTGTGTTTGAATACCTAGATTTGAGATTTATGAGATGTACAGGGGA	25039
Qy	5041	GCTCCAGACTGTGTGGCCCTAAAGTAAAGGCAATCTGATGTGAGATTAAGATTTTGA	5100
Db	25040	GCTCCAGACTGTGTGGCCCTAAAGTAAAGGCAATCTGATGTGAGATTAAGATTTTGA	25099
Qy	5101	AATCATATGAATATACGGTTGTTCAATTAGACACTGTACATGGGTAAAGTATGCTAAAGGAC	5160
Db	25100	AATCATATGAATATACGGTTGTTCAATTAGACACTGTACATGGGTAAAGTATGCTAAAGGAC	25159
Qy	5161	ATGTGTGATGTGATTAACAGAAAGTCAAAGACGGAAACCTTAAGAAATTAATGTATTA	5220
Db	25160	ATGTGTGATGTGATTAACAGAAAGTCAAAGACGGAAACCTTAAGAAATTAATGTATTA	25219
Qy	5221	TTTATTTATTTTATTTATGTTTATTTTAAATTTATTTTATTTATTTATTTATTTTATG	5280
Db	25220	TTTATTTATTTTATTTATGTTTATTTTAAATTTATTTTATTTATTTATTTATTTTATG	25279
Qy	5281	ACGGGAATCTCGCTCTGTGCTCCAGAGCTGAGATGACATGGCGCAACTCAGCTCACTGA	5340
Db	25280	ACGGGAATCTCGCTCTGTGCTCCAGAGCTGAGATGACATGGCGCAACTCAGCTCACTGA	25339
Qy	5341	ACCTCCGCTTCTGTGGTTCAAGGAGAGCTCTGACCTCAAGTATGCTGGAGCTTA	5400
Db	25340	ACCTCCGCTTCTGTGGTTCAAGGAGAGCTCTGACCTCAAGTATGCTGGAGCTTA	25399
Qy	5401	CAGGACCCCAACCACTCACTGACTAATTTTGTATTTTATGAGAGAGGGGTTTCAAC	5460
Db	25400	CAGGACCCCAACCACTCACTGACTAATTTTGTATTTTATGAGAGAGGGGTTTCAAC	25459
Qy	5461	ATGTGTGACAGAGCTGTGCTTGAACCTTGTGACTGTGATTAACCTGTGCTGGCCTTCC	5520
Db	25460	ATGTGTGACAGAGCTGTGCTTGAACCTTGTGACTGTGATTAACCTGTGCTGGCCTTCC	25519
Qy	5521	AAAGTGTGGGATTAACAGTATGACCACTGTGCTGTGGCCTATTTTGTGTTTTTATATAG	5580
Db	25520	AAAGTGTGGGATTAACAGTATGACCACTGTGCTGTGGCCTATTTTGTGTTTTTATATAG	25579
Qy	5581	ATGGAGCTTGTCTATGTTGGCCACAGGCTGCTGAACTCCTGGACTCAAGCAATCTCTCT	5640
Db	25580	ATGGAGCTTGTCTATGTTGGCCACAGGCTGCTGAACTCCTGGACTCAAGCAATCTCTCT	25639
Qy	5641	GCTTGTGCTCTTCAAAAGTTCTGGGATTTACATGTATGATGCTCTGTGCTGTGGCAGAAAT	5700
Db	25640	GCTTGTGCTCTTCAAAAGTTCTGGGATTTACATGTATGATGCTCTGTGCTGTGGCAGAAAT	25699
Qy	5701	CAATATTTTATGATTTTATGTGAAGTAAACCTATGAAAGAAACAGCAAGAGGGGCGAAG	5760
Db	25700	CAATATTTTATGATTTTATGTGAAGTAAACCTATGAAAGAAACAGCAAGAGGGGCGAAG	25759
Qy	5761	AAAAATTAGAGATTTGTGAAACCAAAGAAAGAGATGCTCAGAGAAAGAGCATGTCT	5820
Db	25760	AAAAATTAGAGATTTGTGAAACCAAAGAAAGAGATGCTCAGAGAAAGAGCATGTCT	25819
Qy	5821	ATGATGTCCAAATGCTGTCAAAAGATTAAGAAATTAAGAAATTCATTTGGGTTTATGAAAA	5880
Db	25820	ATGATGTCCAAATGCTGTCAAAAGATTAAGAAATTAAGAAATTCATTTGGGTTTATGAAAA	25879

Qy	5881	GTCTATGGGAAAACCATGATGTAATAAAAAACATTTGTAATGACACAAATGCTTGCAAAAGCATT	5940
Db	25880	GTCTATGGGAAAACCATGATGTAATAAAAAACATTTGTAATGACACAAATGCTTGCAAAAGCATT	25939
Qy	5941	TATAGGGGATCAATTTTGTATTTTCAGAGGACAAACAGTTCATACATAGCGAAGATCTA	6000
Db	25940	TATAGGGGATCAATTTTGTATTTTCAGAGGACAAACAGTTCATACATAGCGAAGATCTA	25999
Qy	6001	GTCTGTGACCAACGGGAGTTAAGTCTGTAAGTGAATGTGAGAGAGAGATCAATTTGAGCTGA	6060
Db	26000	GTCTGTGACCAACGGGAGTTAAGTCTGTAAGTGAATGTGAGAGAGAGATCAATTTGAGCTGA	26059
Qy	6061	GGTGGCTAGACCTGTTCTCATGACACACTAAATGTCATAGATGATCAACAGCTGTGATCCAG	6120
Db	26060	GGTGGCTAGACCTGTTCTCATGACACACTAAATGTCATAGATGATCAACAGCTGTGATCCAG	26119
Qy	6121	TGCCACATCTTTCAGTGAATGACAGAGGGAGTTGAGATTCAGTTGAAATGACCGCTTAAA	6180
Db	26120	TGCCACATCTTTCAGTGAATGACAGAGGGAGTTGAGATTCAGTTGAAATGACCGCTTAAA	26179
Qy	6181	GAAGAGTATGGAAGATGTGG 6201	
Db	26180	GAAGAGTATGGAAGATGTGG 26200	
RESULT 2			
US-10-224-562-3			
Sequence 3, Application US/10224562			
Patent No. 6730506			
GENERAL INFORMATION:			
APPLICANT: YAN, Chunhua et al.			
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC			
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES			
TITLE OF INVENTION: THEREOF			
FILE REFERENCE: CL001098DIY			
CURRENT APPLICATION NUMBER: US/10/224,562			
CURRENT FILING DATE: 2002-08-21			
NUMBER OF SEQ. ID NOS: 10			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 3			
LENGTH: 53332			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-10-224-562-3			
Query Match 100.0%; Score 6201; DB 4; Length 53332;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 6201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	CACCATGAGGCTCAGGCATTTTATGGAAGCCAAATATACGCGCTTAAATAGAAATGTGAC	60
Db	20000	CACCATGAGGCTCAGGCATTTTATGGAAGCCAAATATACGCGCTTAAATAGAAATGTGAC	20059
Qy	61	CTAAATACCCAGAGAAGTCTCCCTTTGTAAATTTGTAAACAAAATTAATATGATGAG	120
Db	20060	CTAAATACCCAGAGAAGTCTCCCTTTGTAAATTTGTAAACAAAATTAATATGATGAG	20119
Qy	121	TTAATAGTTCTATGGAATGATGTAACCCAGAGCCATATCAGCGCTAGCAAAATGTGCGA	180
Db	20120	TTAATAGTTCTATGGAATGATGTAACCCAGAGCCATATCAGCGCTAGCAAAATGTGCGA	20179
Qy	181	ATTATATATATCATCAAAAGTTATCTTCAAGAGCTTCAGCGCTTAATGATGTCTAAAGAA	240
Db	20180	ATTATATATATCATCAAAAGTTATCTTCAAGAGCTTCAGCGCTTAATGATGTCTAAAGAA	20239
Qy	241	ATGTGAAAGCGCTCAGGCATCTGGAAGACAGTGTTAAGCAATTGATCAAAAAGAAAA	300
Db	20240	ATGTGAAAGCGCTCAGGCATCTGGAAGACAGTGTTAAGCAATTGATCAAAAAGAAAA	20299
Qy	301	CCACAGGCGCTTCCCTGCCCTGCCCATCTGATGTAGAGTCTTCATTTCCATGTAGT	360
Db	20300	CCACAGGCGCTTCCCTGCCCTGCCCATCTGATGTAGAGTCTTCATTTCCATGTAGT	20359

```

RESULT 2
US-10-224-562-3
; Sequence 3, Application US/10224562
; Patent No. 6730506
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001098DIV
; CURRENT APPLICATION NUMBER: US/10/224,562
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 53332
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-224-562-3

```

Query Match	100.0%;	Score 6201;	DB 4;	Length 53332;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 6201; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	1	I	ACCACATGGGCTCAGCGCAATTTTGTGAAGCAAAATACAGGCTTTAAATAGATGTGAC	60
Db	2000		CACCATGGGCTCAGCGCAATTTTGTGAAGCCAAATATACAGCTTTAAATAGATGTGAC	20059
Qy	61		CTAAATACCAGAGAACTCCCTTTGTAGATTGTAACTAAATTAATATGAGTAGAG	120
Db	20060		CTAAATATCCAGAGAACTCCCTTTGTAGATTGTAACTAAATTAATATGAGTAGAG	20119
Qy	121		TTAATAGTTCTAATGGAATGTGAACCCAGAGCCATATCAGCGCTAGCCAAATGCGAGA	180
Db	20120		TTAATAGTTCTAATGGAATGTGAACCCAGAGCCATATCAGCGCTAGCCAAATGCGAGA	20179
Qy	181		ATTCAATATATCATCAAAAGTTATCTTCAAGAGCTTCAGCGCTAATGATGCTTAAAGAAA	240
Db	20180		ATTCAATATATCATCAAAAGTTATCTTCAAGAGCTTCAGCGCTAATGATGCTTAAAGAAA	20239
Qy	241		ATGTGAAAGCGCCTCAGGCATCTGAAGGACAGGTTAACAGCAATGATCAAAAGAAAAA	300
Db	20240		ATGTGAAAGCGCCTCAGGCATCTGAAGGACAGGTTAACAGCAATGATCAAAAGAAAAA	20239
Qy	301		CCACAGGCGCTTCCCTTCGCCCAATCTTGATGTAGCAGTCTTCATTTCCATGTAGT	360
Db	20300		CCACAGGCGCTTCCCTTCCGCCCAATCTTGATGTAGCAGTCTTCATTTCCATGTAGT	20359

QY 361 AAATTTCTAGATACAGCTTGTAGAGCTCAAACTGGAAGAAAGCTCCATTCAG 420
DB 20360 AAAATTTCTAGATACAGCTTGTAGAGCTCAAACTGGAAGAAAGCTCCATTCAG 20419
QY 421 GAAATTTCTAGATACAGCTTGTAGAGCTCAAACTGGAAGAAAGCTCCATTCAG 480
DB 20420 GAAATTTCTAGATACAGCTTGTAGAGCTCAAACTGGAAGAAAGCTCCATTCAG 20479
QY 481 TGTAGCCGCGCGCGCTGCTCAAGCTGTATCCAGCCCTTTGGAGGCCAGATGGG 540
DB 20480 TGTAGCCGCGCGCGCTGCTCAAGCTGTATCCAGCCCTTTGGAGGCCAGATGGG 20539
QY 541 CAGATCAGAGGTCAAGAGTTTGAACAAGCTTACCAACATGTGAAACCCCGTCTTA 600
DB 20540 CAGATCAGAGGTCAAGAGTTTGAACAAGCTTACCAACATGTGAAACCCCGTCTTA 20599
QY 601 CTAAAGATCAAAAAATTAGCCAGGTGTGGCGCAACCTGTAAACCCAGCTGTCA 660
DB 20600 CTAAAGATCAAAAAATTAGCCAGGTGTGGCGCAACCTGTAAACCCAGCTGTCA 20659
QY 661 GAGAGTGAAGCAGAGAAATGCTTGAACCCAGAGGCCAGAGGTGACAGCAGCAAGATCA 720
DB 20660 GAGAGTGAAGCAGAGAAATGCTTGAACCCAGAGGCCAGAGGTGACAGCAGCAAGATCA 20719
QY 721 CACCAATGCACTGTAGCTGTAGATGACAGGCGCAAGCTTCACTCAAAAAA 780
DB 20720 CACCAATGCACTGTAGCTGTAGATGACAGGCGCAAGCTTCACTCAAAAAA 20779
QY 781 AAGAAATATGTAAGTGTGTCTAATAAATTAAGCAAGTGAAGCAAGCTTAAG 840
DB 20780 AAGAAATATGTAAGTGTGTCTAATAAATTAAGCAAGTGAAGCAAGCTTAAG 20839
QY 841 CCTATGACCATGCTAAGTAAGTAAGTGTGGAACAATTAAGGAACTTAACCAT 900
DB 20840 CCTATGACCATGCTAAGTAAGTAAGTGTGGAACAATTAAGGAACTTAACCAT 20899
QY 901 CCTGGAAGTAAGGTTTGGAAAAGAAATGTTTGAAGCAAAAGGTTAAAGAGTGA 960
DB 20900 CCTGGAAGTAAGGTTTGGAAAAGAAATGTTTGAAGCAAAAGGTTAAAGAGTGA 20959
QY 961 AAAAAATTAATTAACCACTTGTGTGTGGAAGAAATGGGATAGGAGCTTAAGAAAT 1020
DB 20960 AAAAAATTAATTAACCACTTGTGTGTGGAAGAAATGGGATAGGAGCTTAAGAAAT 21019
QY 1021 CAAATAGAAATGTTTCAATGTATTAAGAACCTGTGAAGGTTGAAGCAATTAATTA 1080
DB 21020 CAAATAGAAATGTTTCAATGTATTAAGAACCTGTGAAGGTTGAAGCAATTAATTA 21079
QY 1081 TGTGACCATGCGCGGACTTTTATGTATGTAATGCTTGGCAATTTAAATAGAGAGA 1140
DB 21080 TGTGACCATGCGCGGACTTTTATGTATGTAATGCTTGGCAATTTAAATAGAGAGA 21139
QY 1141 GAGAAATGTAGACAGTTGGAATGAGTGAAGTTCTGCCAGCAATGTGAAGAAAG 1200
DB 21140 GAGAAATGTAGACAGTTGGAATGAGTGAAGTTCTGCCAGCAATGTGAAGAAAG 21199
QY 1201 AAGACAGTGAAGCAGAGTTTGAAGATTAATCAAGACAGAAATTAATGTGCTGCGCAGT 1260
DB 21200 AAGACAGTGAAGCAGAGTTTGAAGATTAATCAAGACAGAAATTAATGTGCTGCGCAGT 21259
QY 1261 GGCATCTAGTCTGAGTCTTAATCTGAGGAAAGAAAGTGAAGATTAAGAGCTTGTAGT 1320
DB 21260 GGCATCTAGTCTGAGTCTTAATCTGAGGAAAGAAAGTGAAGATTAAGAGCTTGTAGT 21319
QY 1321 TAGAAGAGAGTGAAGGCTTCAAGGACCTACAGGTGTGATTAATTAAGAAATGATG 1380
DB 21320 TAGAAGAGAGTGAAGGCTTCAAGGACCTACAGGTGTGATTAATTAAGAAATGATG 21379
QY 1381 GAGAAAGAAATTAATCTGAGAGAGTGAATTTTCAAGCTTGAAGTCTTCAATACAG 1440
DB 21380 GAGAAAGAAATTAATCTGAGAGAGTGAATTTTCAAGCTTGAAGTCTTCAATACAG 21439

QY 1441 CACTGTCTAAATGCTTCAAAAGCATGATCCCTGCCCTCAAGGAGCTTACAGCCAAAAC 1500
DB 21440 CACTGTCTAAATGCTTCAAAAGCATGATCCCTGCCCTCAAGGAGCTTACAGCCAAAAC 21499
QY 1501 AAGAGATGAAGATACACACCAATACTATTATAGACACTTGTTGTAATTAAGAAAG 1560
DB 21500 AAGAGATGAAGATACACACCAATACTATTATAGACACTTGTTGTAATTAAGAAAG 21559
QY 1561 AATATGATCTAGTACTGTAGATGTGCAACGCGCATCAAGATATCTTCTAGTTTCAAGA 1620
DB 21560 AATATGATCTAGTACTGTAGATGTGCAACGCGCATCAAGATATCTTCTAGTTTCAAGA 21619
QY 1621 GTTTCAAGATCGGCGCGGCTGCTCAAGCTGTATCCAGACCTTTGGAGGCCA 1680
DB 21620 GTTTCAAGATCGGCGCGGCTGCTCAAGCTGTATCCAGACCTTTGGAGGCCA 21679
QY 1681 GCGCGGTGATCAAAAGTCAAGAGTCAAGACATCTGTGTTAAACGCGTGAACCCCG 1740
DB 21680 GCGCGGTGATCAAAAGTCAAGAGTCAAGACATCTGTGTTAAACGCGTGAACCCCG 21739
QY 1741 TCTCTACAAAAATATTAATAAATTAGCCAGGCGTGTGGCGCGCTGTAGTCCAGCT 1800
DB 21740 TCTCTACAAAAATATTAATAAATTAGCCAGGCGTGTGGCGCGCTGTAGTCCAGCT 21799
QY 1801 ACTCAGAGGCTGAGCAGAGAAATGCGTGAACCCGGAAGTGAAGTTTGGTGAAGCCG 1860
DB 21800 ACTCAGAGGCTGAGCAGAGAAATGCGTGAACCCGGAAGTGAAGTTTGGTGAAGCCG 21859
QY 1861 AGATCGCGCCTGAGCTCAGCTGTGGGCGACAGAGTGAAGTGTCTCAAAAAA 1920
DB 21860 AGATCGCGCCTGAGCTCAGCTGTGGGCGACAGAGTGAAGTGTCTCTCAAAAAA 21919
QY 1921 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1980
DB 21920 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 21979
QY 1981 AGCATGTTTCAAGGAGAGAAATGTTTCAAGAAAGTTCAGAAAGTTCAGAAAGTTCAG 2040
DB 21980 AGCATGTTTCAAGGAGAGAAATGTTTCAAGAAAGTTCAGAAAGTTCAGAAAGTTCAG 22039
QY 2041 ATGTTCAAGAGAGAGACCCAGAGTCAATGTTTGTAGGTTAAGAGAAACAGTGT 2100
DB 22040 ATGTTCAAGAGAGAGACCCAGAGTCAATGTTTGTAGGTTAAGAGAAACAGTGT 22099
QY 2101 TTGCAATCTCAGGTTTCAATGAGTGTGTAATTAATGAGTGTGAGCACTGCA 2160
DB 22100 TTGCAATCTCAGGTTTCAATGAGTGTGTAATTAATGAGTGTGAGCACTGCA 22159
QY 2161 TTTTAAAAATGAATTAATGAGTGAAGAAATTAATTAATTAATTAATTAATTAATTA 2220
DB 22160 TTTTAAAAATGAATTAATGAGTGAAGAAATTAATTAATTAATTAATTAATTAATTA 22219
QY 2221 TTGAAGAGAGATTAATTTCTGCACTTTTGTGCAATTTGTAATGTAATTAATTAAT 2280
DB 22220 TTGAAGAGAGATTAATTTCTGCACTTTTGTGCAATTTTGTGCAATTTTGTGCAATTT 22279
QY 2281 TTATGTATGATGTGAATTAACGATATACATATTAATTTCTTACTGTAGACTGCA 2340
DB 22280 TTATGTATGATGTGAATTAACGATATACATATTAATTTCTTACTGTAGACTGCA 22339
QY 2341 AATCTTTAAAGCATGCGCTGTGCTTAATCTTCTTATTTTGCAGAGAGAAATCCAGATC 2400
DB 22340 AATCTTTAAAGCATGCGCTGTGCTTAATCTTCTTATTTTGCAGAGAGAAATCCAGATC 22399
QY 2401 TGAGAGAGCAAACTTTTGTGAGGTTTGAAGCAAGCTTATGCAATTTGCTAAAGTGA 2460
DB 22400 TGAGAGAGCAAACTTTTGTGAGGTTTGAAGCAAGCTTATGCAATTTGCTAAAGTGA 22459
QY 2461 TTCTTAGTTAAATTTCTTCCCATAGTGCATTAATCTTCTAGTTCTGTGGCTGA 2520
DB 22460 TTCTTAGTTAAATTTCTTCCCATAGTGCATTAATCTTCTAGTTCTGTGGCTGA 22519
QY 2521 AATTAAGAAATATTAATTAAGAAACAGCATACAAAGTCTGGGAAATTAATTTGGGTAG 2580

Db 22520 AATACGATATATATGTAAGACATACAAAGCTGGGAAATATATGAGTGGTG 22579
QY 2581 GCTGAGAGCCTCATTTCTAAGAAATGTGACCTTAGCAGGGATAGTGGCTCACACT 2640
Db 22580 GCTGAGAGCCTCATTTCTAAGAAATGTGACCTTAGCAGGGATAGTGGCTCACACT 22639
QY 2641 ATTAATTCAGACATTTGGGAGGCGCAAGTCAAGAAATCGTTGAACCCAAAGATTCAAGA 2700
Db 22640 ATTAATTCAGACATTTGGGAGGCGCAAGTCAAGAAATCGTTGAACCCAAAGATTCAAGA 22699
QY 2701 CTAGCATGGGCAACATAGCAAGACCTCATCTACAAAAAATTAAAAATCAGCTGAGCA 2760
Db 22700 CTAGCATGGGCAACATAGCAAGACCTCATCTACAAAAAATTAAAAATCAGCTGAGCA 22759
QY 2761 TGGTGGCATACGCTGTAGTCCCACTACCTGGAAAGTATGATGGGTGATCGCTTGACA 2820
Db 22760 TGGTGGCATACGCTGTAGTCCCACTACCTGGAAAGTATGATGGGTGATCGCTTGACA 22819
QY 2821 CAGGAGTTTGAAGCTAAGGTGAGCCATGATCACAACATGCACTCCAGCTTGATGACAG 2880
Db 22820 CAGGAGTTTGAAGCTAAGGTGAGCCATGATCACAACATGCACTCCAGCTTGATGACAG 22879
QY 2881 AGGAAGACCCCTGCTCCCTAAAAAAGAAATGTGATTTTATCTTGAACAGTACAGT 2940
Db 22880 AGGAAGACCCCTGCTCCCTAAAAAAGAAATGTGATTTTATCTTGAACAGTACAGT 22939
QY 2941 CATTAAGTCAATTAAGTTGATGAGAAATATATATATGATCAGAGAAATTTATATCACT 3000
Db 22940 CATTAAGTCAATTAAGTTGATGAGAAATATATATATGATCAGAGAAATTTATATCACT 22999
QY 3001 GGGGTGTAGAGATATATGAAGAATAAGAACTAGAGTCAAGGATTCACATTAAAGT 3060
Db 23000 GGGGTGTAGAGATATATGAAGAATAAGAACTAGAGTCAAGGATTCACATTAAAGT 23059
QY 3061 TTTGTTGTTTGTGTTTGAAGCAGAGTCTCTTTTGTTAACCAAGCTAGAGTCAATGATG 3120
Db 23060 TTTGTTGTTTGTGTTTGAAGCAGAGTCTCTTTTGTTAACCAAGCTAGAGTCAATGATG 23119
QY 3121 CAGTCAATGCTCACCGCAGCTCAAACTCCCAAGCTCAAAATATCTTCCAGCTGGGCTT 3180
Db 23120 CAGTCAATGCTCACCGCAGCTCAAACTCCCAAGCTCAAAATATCTTCCAGCTGGGCTT 23179
QY 3181 CCCAAGTCTGGAATTAGAGGTGAGCCAAAGGTTTATGATTTGATCTGGCTTAAT 3240
Db 23180 CCCAAGTCTGGAATTAGAGGTGAGCCAAAGGTTTATGATTTGATCTGGCTTAAT 23239
QY 3241 GCCTCTCAAACTTCAGTGAAGCAGCAAGTGAACGGGAACTGATCAACAAGCTGGGTT 3300
Db 23240 GCCTCTCAAACTTCAGTGAAGCAGCAAGTGAACGGGAACTGATCAACAAGCTGGGTT 23299
QY 3301 TAAAGCTGAGCCTCTGCAATTTCAATGATCAAGCTGATGCAATGATGCTGGTCAAGA 3360
Db 23300 TAAAGCTGAGCCTCTGCAATTTCAATGATCAAGCTGATGCAATGATGCTGGTCAAGA 23359
QY 3361 GCCAAGCATGAGCAGCAAGATCTAGTTAGCAATTAGTAATCAAGGTGATTAATGAT 3420
Db 23360 GCCAAGCATGAGCAGCAAGATCTAGTTAGCAATTAGTAATCAAGGTGATTAATGAT 23419
QY 3421 AGTGAATTAAGATGAGAGAAATGTGAATTCAGTAAACAAGAGATTCACTCTTG 3480
Db 23420 AGTGAATTAAGATGAGAGAAATGTGAATTCAGTAAACAAGAGATTCACTCTTG 23479
QY 3481 GTTAATCTGAGCATGAGAGAGAAAGATGAGGCGCAAACTAACTGGTTTGTGTTGACT 3540
Db 23480 GTTAATCTGAGCATGAGAGAGAAAGATGAGGCGCAAACTAACTGGTTTGTGTTGACT 23539
QY 3541 GACGAGAGAAATTTGATGCTCTATTAACAAGAAATAGGAGAAAGTTGTTGAGAGAAA 3600
Db 23540 GACGAGAGAAATTTGATGCTCTATTAACAAGAAATAGGAGAAAGTTGTTGAGAGAAA 23599
QY 3601 GAGAGTCTGTGTTTGAAGCTGTGTTGAGTCCAGGTGAGACAGATCTCCAAAGGAAAT 3660
|||||

Db 23600 GAGAGTCTGTGTTTGAAGCTGTGTTGAGTCCAGGTGAGACAGATCTCCAAAGGAAAT 23659
QY 3661 GAGCAGTAGCAACCTAAAGAGAAATCTGTGCTCAAGAGGAGCTGTGACTGAGTGT 3720
Db 23660 GAGCAGTAGCAACCTAAAGAGAAATCTGTGCTCAAGAGGAGCTGTGACTGAGTGT 23719
QY 3721 AGATCTGAGGGTCACTCAGCACTAGAGTTTAAAGACAAAGAGTAGGCAACCAAGAGAC 3780
Db 23720 AGATCTGAGGGTCACTCAGCACTAGAGTTTAAAGACAAAGAGTAGGCAACCAAGAGAC 23779
QY 3781 AAATACACAAGAGAGAGAGAGCTGATGATGAGACTTTTCCCTTTTATGATGAGAGAGGA 3840
Db 23780 AAATACACAAGAGAGAGAGAGCTGATGATGAGACTTTTCCCTTTTATGATGAGAGAGGA 23839
QY 3841 ACAGAAATGAAGAGATGAAGGAGAGCAGCTTGAGAAATGTAGAGCATCTGAAAAAAA 3900
Db 23840 ACAGAAATGAAGAGATGAAGGAGAGCAGCTTGAGAAATGTAGAGCATCTGAAAAAAA 23899
QY 3901 ATACACACTGTATGAGAAATCAAGGAGAGAAATTTCAAGAGAGGATATGTGAGCA 3960
Db 23900 ATACACACTGTATGAGAAATCAAGGAGAGAAATTTCAAGAGAGGATATGTGAGCA 23959
QY 3961 GTATTACAAGCATCAGAAATACAGCTAAAGTCTATCTTGAATGCAATTGACTTGTGG 4020
Db 23960 GTATTACAAGCATCAGAAATACAGCTAAAGTCTATCTTGAATGCAATTGACTTGTGG 24019
QY 4021 ATTTGTGAGGACACACTAATTAATTAAGAAATTTATGTGGGTATATGAGGCAAAAG 4080
Db 24020 ATTTGTGAGGACACACTAATTAATTAAGAAATTTATGTGGGTATATGAGGCAAAAG 24079
QY 4081 GAAAGGTTATCCAAATCAAGACAGGTGGAGATGAGTCTCCAAAGGTGAGAGCA 4140
Db 24080 GAAAGGTTATCCAAATCAAGACAGGTGGAGATGAGTCTCCAAAGGTGAGAGCA 24139
QY 4141 TCAGTGAATGTGGAGAGGCGCAGAGCATCCATGCCATCCAGGCAAGCCACTCCA 4200
Db 24140 TCAGTGAATGTGGAGAGGCGCAGAGCATCCATGCCATCCAGGCAAGCCACTCCA 24199
QY 4201 GAAAGCTCCATGAGAGTTCAAGTATCCAGAGGTCTGTATCCCTAATCTTTCTGGGTTT 4260
Db 24200 GAAAGCTCCATGAGAGTTCAAGTATCCAGAGGTCTGTATCCCTAATCTTTCTGGGTTT 24259
QY 4261 TGCATAGGCTCAATGTGAGGAGCATGATTTAATCAATGAGGCACTGGTGAATCAACT 4320
Db 24260 TGCATAGGCTCAATGTGAGGAGCATGATTTAATCAATGAGGCACTGGTGAATCAACT 24319
QY 4321 AACCTTCAACCCCTCTCCCTCCCTAATCATGCTTGTGCTTTCCAGTGAACCAATCCCTA 4380
Db 24320 AACCTTCAACCCCTCTCCCTCCCTAATCATGCTTGTGCTTTCCAGTGAACCAATCCCTA 24379
QY 4381 TCTTAAGCTACCCCAATGCTTGGCCAGCTATCAGTCACTTACAAAAAGCATCACTTGG 4440
Db 24380 TCTTAAGCTACCCCAATGCTTGGCCAGCTATCAGTCACTTACAAAAAGCATCACTTGG 24439
QY 4441 GAGATTTCAAGGATTTTGAAGTTGGCTGTGAGAAATTTAGTTGAAGTCAAAATTAAT 4500
Db 24440 GAGATTTCAAGGATTTTGAAGTTGGCTGTGAGAAATTTAGTTGAAGTCAAAATTAAT 24499
QY 4501 TTCACAATATCAACAGTGTGTATTTTATATCAGGCGCATTTAAATGTTTTTAAACAAAG 4560
Db 24500 TTCACAATATCAACAGTGTGTATTTTATATCAGGCGCATTTAAATGTTTTTAAACAAAG 24559
QY 4561 AGGTATTAATTTCAAGATTTTCTTTTATTAAGCTTACCTGATGACAGTGTGTGAATAG 4620
Db 24560 AGGTATTAATTTCAAGATTTTCTTTTATTAAGCTTACCTGATGACAGTGTGTGAATAG 24619
QY 4621 ATTTGGAATGAGGCAATCTTTTATTTTGAAGTATTAATCCCTGACCTTACTTCTC 4680
Db 24620 ATTTGGAATGAGGCAATCTTTTATTTTGAAGTATTAATCCCTGACCTTACTTCTC 24679
QY 4681 CTGTTTTCTTCTACCTCTCTCCCTCACTCAACAGAAAACTTCTCTCCCTCACTCAT 4740
Db 24680 CTGTTTTCTTCTACCTCTCTCCCTCACTCAACAGAAAACTTCTCTCCCTCACTCAT 24739
|||||

QY 4741 TCCCTGAATGCTGCTGCTGTTAAGGTTCCAGCTTGAAGAGGCTTAATCAAGAACAC 4800
 DB 24740 TCCCTGAAGGCTGCTGCTGTTAAGGTTCCAGCTTGAAGAGGCTTAATCAAGAACAC 24799
 QY 4801 AGTGGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4860
 DB 24800 AGTGGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24859
 QY 4861 GCCGGAATGACGGGGAAGAAACAAGAGTCAAGCAATATTTTCTGAGTCTGAGTATTTAA 4920
 DB 24860 GCCGGAATGACGGGGAAGAAACAAGAGTCAAGCAATATTTTCTGAGTCTGAGTATTTAA 24919
 QY 4921 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4980
 DB 24920 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24979
 QY 4981 ATACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5040
 DB 24980 ATACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25039
 QY 5041 GCTCCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5100
 DB 25040 GCTCCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25099
 QY 5101 AATCATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5160
 DB 25100 AATCATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25159
 QY 5161 ATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5220
 DB 25160 ATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25219
 QY 5221 TTTATATTTTAT 5280
 DB 25220 TTTATATTTTAT 25279
 QY 5281 AGGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5340
 DB 25280 AGGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25339
 QY 5341 AACTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5400
 DB 25340 AACTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25399
 QY 5401 CAGGCAACCAACCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 5460
 DB 25400 CAGGCAACCAACCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 25459
 QY 5461 ATGTGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5520
 DB 25460 ATGTGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25519
 QY 5521 AAAAGTCTGGGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 5580
 DB 25520 AAAAGTCTGGGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 25579
 QY 5581 ATGGGGTCTTGTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5640
 DB 25580 ATGGGGTCTTGTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25639
 QY 5641 GCTTGGCTCTCAAAAGTTCTGGGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 5700
 DB 25640 GCTTGGCTCTCAAAAGTTCTGGGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 25699
 QY 5701 CAATATATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 5760
 DB 25700 CAATATATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 25759
 QY 5761 AAAAATTAGAGATTGTGGAACCAAGAAAGAGAGTGTCTCAGAGAAAGAGAGATGTGTCT 5820
 DB 25760 AAAAATTAGAGATTGTGGAACCAAGAAAGAGAGTGTCTCAGAGAAAGAGAGATGTGTCT 25819

QY 5821 ATGATGCCAATGCTGCAAGATTAAGAAATTAAGATTAAGATTAAGATTAAGATTAAG 5880
 DB 25820 ATGATGCCAATGCTGCAAGATTAAGAAATTAAGATTAAGATTAAGATTAAGATTAAG 25879
 QY 5881 GTCATGGGAAACCATGCTGTAAGAAACATTTGTGATGACAAATCGTTGCAAAAGCATTTT 5940
 DB 25880 GTCATGGGAAACCATGCTGTAAGAAACATTTGTGATGACAAATCGTTGCAAAAGCATTTT 25939
 QY 5941 TATAGGGGATTAATTTGTATTTTCAAGAGCAAAAGTTCATATCAATGCAATGCTTA 6000
 DB 25940 TATAGGGGATTAATTTGTATTTTCAAGAGCAAAAGTTCATATCAATGCAATGCTTA 25999
 QY 6001 GTGTGACCAAGGAGATTAAGTCTGAAGTGAATTTGGAAGAGAGATGATGAGCTGA 6060
 DB 26000 GTGTGACCAAGGAGATTAAGTCTGAAGTGAATTTGGAAGAGAGATGATGAGCTGA 26059
 QY 6061 GGTGGCTGAGCTGCTTCTCATGACACTAATGTCATGAGTCAACAGCTGTGATCAAG 6120
 DB 26060 GGTGGCTGAGCTGCTTCTCATGACACTAATGTCATGAGTCAACAGCTGTGATCAAG 26119
 QY 6121 TSCCAATCTTCACTGAATGACAGAGAGATTAAGATTAAGATTAAGATTAAGATTAAG 6180
 DB 26120 TSCCAATCTTCACTGAATGACAGAGAGATTAAGATTAAGATTAAGATTAAGATTAAG 26179
 QY 6181 GAAGATTAATGGAAGATGCTG 6201
 DB 26180 GAAGATTAATGGAAGATGCTG 26200

RESULT 3
 US-08-544-900-1
 ; Sequence 1, Application US/08544900
 ; Patent No. 5736331
 ; GENERAL INFORMATION:
 ; APPLICANT: Lin, Stanley Li
 ; TITLE OF INVENTION: Method for Identifying Nucleic
 ; TITLE OF INVENTION: Acids Encoding c-fos Promoter
 ; NUMBER OF INVENTION: Activating Proteins
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Schering-Plough Corporation
 ; ADDRESS: Patent Department K-6-1 (1990)
 ; STREET: 2000 Galloping Hill Road
 ; CITY: Kenilworth
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07033-0530
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Apple Macintosh
 ; OPERATING SYSTEM: Macintosh 7.1
 ; SOFTWARE: Microsoft Word 5.1a
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/544,900
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/272,412
 ; FILING DATE: 8-JUL-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Thompson, Paul A.
 ; REGISTRATION NUMBER: 35,385
 ; REFERENCE/DOCKET NUMBER: OCO439K
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 908 298 5150
 ; TELEFAX: 908 298 5388
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1930 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-544-900-1

Query Match 6.5%; Score 402.6; DB 1; Length 1930;
Best Local Similarity 92.5%; Pred. No. 1,1e-85;
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

QY 16 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGAGACCTAAATCCAGAG 75
DB 245 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGAGACCTAAATCCAGAG 304
QY 76 AACTCCCTTTGTAAGTTGTACAAAATTAATAGAGAGTTAAATAGTTCTAAG 135
DB 305 CACCCCTTTGTAAGTTGTACAAAATTAATAGAGAGTTAAATAGTTCTAAG 364
QY 136 GAATGTGAACCCAGAGCCATATCAGCGTAGCAAAATGCGAAATTCATATCATCA 195
DB 365 GAGTGTGAGACCCAGAGCCATATCAGCGTAGCAAAATGCGAAATTCATATCATCA 424
QY 196 AAGTTATCTTCAAGAGCTTCAGCGCTTAATGATGCTTAAAGAAATGTGAACGCGCTC 255
DB 425 AAGTTATCTTCAAGAGCTTCAGCGCTTAATGATGCTTAAAGAAATGTGAACGCGCTC 484
QY 256 AGCCATCTGAAGACAGTGTTAACAGCAATTGATCAAAAAGAAACCCAGCGCCCTCC 315
DB 485 AGCCGCGCGAGAGACAGTGTTAACAGCAATTGATCAAAAAGAAACCCAGCGCCCTCC 544
QY 316 CTTCCCTCACTGATGATGAGAGAGTCTTCATTTCCATAGTAAATTTCTAGATAC 375
DB 545 CTTCCCTCACTGATGATGAGAGAGTCTTCATTTCCATAGTAAATTTCTAGATAC 604
QY 376 AGCTGTAGAGCTCAAGTACTGAAAGAAAGTCCCATCAAGAAATTTATCTTAAG 435
DB 605 GTCTGTAGAGCTCAAGTACTGAAAGAAAGTCCCATCAAGAAATTTATCTTAAG 664
QY 436 ATACTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 483
DB 665 ATACTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 713

RESULT 4
PCT-US95-07874-1
Sequence 1, Application PC/TUS9507874
GENERAL INFORMATION:

APPLICANT:
TITLE OR INVENTION: Method for Identifying Nucleic Acids Encoding c-fos Promoter A
NUMBER OF SEQUENCES: 2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07874
FILING DATE: -June-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/272,412
FILING DATE: 8-JUL-1994
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1930 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
PCT-US95-07874-1

Query Match 6.5%; Score 402.6; DB 5; Length 1930;
Best Local Similarity 92.5%; Pred. No. 1,1e-85;
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;
QY 16 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGAGACCTAAATCCAGAG 75

DB 245 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGAGACCTAAATCCAGAG 304
QY 76 AACTCCCTTTGTAAGTTGTACAAAATTAATAGAGAGTTAAATAGTTCTAAG 135
DB 305 CACCCCTTTGTAAGTTGTACAAAATTAATAGAGAGTTAAATAGTTCTAAG 364
QY 136 GAATGTGAACCCAGAGCCATATCAGCGTAGCAAAATGCGAAATTCATATCATCA 195
DB 365 GAGTGTGAGACCCAGAGCCATATCAGCGTAGCAAAATGCGAAATTCATATCATCA 424
QY 196 AAGTTATCTTCAAGAGCTTCAGCGCTTAATGATGCTTAAAGAAATGTGAACGCGCTC 255
DB 425 AAGTTATCTTCAAGAGCTTCAGCGCTTAATGATGCTTAAAGAAATGTGAACGCGCTC 484
QY 256 AGCCATCTGAAGACAGTGTTAACAGCAATTGATCAAAAAGAAACCCAGCGCCCTCC 315
DB 485 AGCCGCGCGAGAGACAGTGTTAACAGCAATTGATCAAAAAGAAACCCAGCGCCCTCC 544
QY 316 CTTCCCTCACTGATGATGAGAGAGTCTTCATTTCCATAGTAAATTTCTAGATAC 375
DB 545 CTTCCCTCACTGATGATGAGAGAGTCTTCATTTCCATAGTAAATTTCTAGATAC 604
QY 376 AGCTGTAGAGCTCAAGTACTGAAAGAAAGTCCCATCAAGAAATTTATCTTAAG 435
DB 605 GTCTGTAGAGCTCAAGTACTGAAAGAAAGTCCCATCAAGAAATTTATCTTAAG 664
QY 436 ATACTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 483
DB 665 ATACTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 713

RESULT 5
US-09-497-855A-38/c
Sequence 38, Application US/09497855A
Patent No. 6605432
GENERAL INFORMATION:
APPLICANT: Huang, Tim
TITLE OR INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
FILE REFERENCE: UMO1523
CURRENT APPLICATION NUMBER: US/09/497,855A
CURRENT FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/120,592
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: 60/118,760
PRIOR FILING DATE: 1999-02-05
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn version 3.0
SEQ ID NO 38
LENGTH: 128779
TYPE: DNA
ORGANISM: Homo sapiens;
US-09-497-855A-38

Query Match 5.3%; Score 327.6; DB 4; Length 128779;
Best Local Similarity 54.4%; Pred. No. 5,2e-67;
Matches 808; Conservative 0; Mismatches 664; Indels 12; Gaps 7;

QY 459 TGTACATTGGAATATATATAGTTGTTAGCTGGCGCGGTGCTCAAGCTGTAAATCCAG 518
DB 3690 TGTATCTAAAT 36841
QY 519 CCGTTGGGAGGCGAGAGTGGCGAGATCATAGAGTCCGAGATTTGAGACCGAGCTTACCA 578
DB 36840 CACTTGGGAGGCGAGAGTGGCGAGATCATAGAGTCCGAGATTTGAGACCGAGCTTACCA 36781
QY 579 ACATGTGAACCCCGTCTCTACTTAAGATACAAAATTTAGCCAGGTGTGTGCGCAC 638
DB 36780 ACAGGTGAACCCCGTCTCTACTTAAGATACAAAATTTAGCCAGGTGTGTGCGCGGC 36721
QY 639 ACCTGTAAACCCAGTGTCTGAGAGAGTGAAGGCGAGAGATTTGTTGAATCCAGAGGCA 698
DB 36720 ACCTGTATCTCCAGTACTCTGCGAGGCTGAGGCGAGAGATTTGTTGAATCCAGAGGCA 36661

Db 12803 TTAAAGCTGAGTGAAGGCGCGGCGGCTGAGCTCAGCTGTAATCCAGCACTTTGAG 12862
QY 1674 AGGCCGAGCGGTGATCACAAGTCAAGAGATCAAGACCATCTGTGTTAACAGGTGA 1733
Db 12863 AGGTGGAGGTGGGTGATCAAGAGTCAAGAGATCCAGCATCTGTGCTAACAGGTGA 12922
QY 1734 AACCCGCTCTCAAAAAATATATAAATTAAGCCAGGCGGTGAGCGGCGGCTGTAGT 1793
Db 12923 AACCCGCTCTCAAAAAATATATAAATTAAGCGGTGAGGTGAGCGGCGGCTGTAGT 12982
QY 1794 CCCAGCTACTCAGAGGCTGAGGCGAGAGATGCGGTGAACCCGAGAGTGAAGTTGC- 1852
Db 12983 CCCAGCTACTCAGAGGCTGAGGCGAGAGATGCGGTGAACCCGAGAGTGAAGTTGCA 13042
QY 1853 GTGAGCGGAGATGCGGCGCATCTGCGCTCAGACCTGGGCGAGAGTGAAGTCTGCA 1912
Db 13043 GTGAGCTGAGCTGCGACCACTGCACTCAGCTGGGCGAGAGTGAAGTCTGCTCAA 13102
QY 1913 AAAAAAAAAAAAAAAAAAGTTTCAGATCTTAACACATGCACTTCAACAGTCTAG 1972
Db 13103 AAAAAAAAAAAAAAAAAAGTTTCAGATCTTAACTGATCTTTTAACTTTCTTTT 13162
QY 1973 AATAGAGAGCATGTTTACAGGAGAGAAAAATGTTTTCAGAAAGTACAGATGAGGAAA 2032
Db 13163 TTAGAGACAGAGTCTCATCTCATCACCCATGCTGAGATACAGTGTGCCATCTGTCTCA 13222
QY 2033 TAAAGATATGTTTCAAGAAAGAGACCCAGAGTCAATGTTGT-----TAGGTTAG 2085
Db 13223 CTCGAATCTTGGGCTCTGAGTTCAACCAATCTCAGCTCCCAATACCTGG 13282
QY 2086 AGGAAACACAGTGTGTTTCAATCTCAGGTTCCATTAGTGCATTAGTAATCAATAGT 2145
Db 13283 GACCAAGGAGAGCTGCCACCAAGCTAAATTTTGGGTATTTTGTAGAGATGGG 13342
QY 2146 GGTTCACAACTGCAATTTTAAAAATGAATTAATGATGAGAGAAATGAATATTT 2205
Db 13343 GCCTCACAGTGTGCTCAGGTGTGCTGAATCCCTGAGCTCAAGATCACTTCCTC 13402
QY 2206 AGCATGATTAATCTTTGAAAAGCAAGTATTTTCTGCAACTTTGCTCAATTTGA 2265
Db 13403 GGCCTGC-CAAAGTGTGGGATTAATGAGCAATGAGCACTGACCTCCCAATTTTAA 13461
QY 2266 ACTGACTTATATTTTATGATGATGTAATACAGATACATATATATTTCTACTGT 2325
Db 13462 TATTTATATTTATTTTATTTTACTTATTTATTTTGAAGAGAGGTCTACTCTGCACC 13521
QY 2326 AGACTGCACTCAAAAAATCTTTAAGCACTGCGCTGCTTAACTTCTTATTTTGCAGAG 2385
Db 13522 CAGGCTGAGATACAGTGGCACTATCTCAGCTCACTGCAACCTTGCCTCGGTTCAAG 13581
QY 2386 GAGAAATCCAAAGTCTGAGAGGACAAACATTTTCTGAGGTATAGAAACAGCTTATGC 2445
Db 13582 CGAATCTGCTGCTCAGCTCTGAGTACCTGGATTAACAGGATCACCAACATCCCC 13641
QY 2446 CATTGCTAAAGATTTCTTATGTTAAATTTCTTCCCACTAGTGCATGCACTTCTA 2505
Db 13642 GTTAAATTTTGTATTTTATTTTATTTAGTAGAGAGGTTTC-----ACGTTGTGCCAGAT 13694
QY 2506 GTTCTGTGCTGCAATAACAGATATATTTAGTAAACAGCATACAAATCTGGGAAA 2565
Db 13695 GGTCTGCAACTCTGACCTCAAGTATATCCCACTCAGCTCCCAAGTGTGAG--- 13751
QY 2566 TATATTGGGTAGTGTGCTGAGAGCTCATTTTCTAAGAAATGTGACCTTAGCAGGTA 2625
Db 13752 -ATTATAGGTGAGCACTCGCTGATGTTTAAAAAGTGGGTCAATGGGCTTGGGG 13810
QY 2626 TGTGCTGCACTCACTTAATTTCAAGCATTTGGGAGGCAAGTCAAGATGCTTGA 2685
Db 13811 CGGTGCTCACTGCTTAAATCCAGCATTTTGTATACGAGGCGGAGTGAATCAAGGT 13870
QY 2686 CCCAAGAGTTCAAGATGAGTGGGCAATAGCAAGACCTCATCTCAAAAAAT-TT 2744

Db 13871 --CAGAGATCGAGACCATCTGCTTAACAGGTGAACCCGCTCTACTAAAAATACA 13928
QY 2745 AAAATCAGCTGAGCATGTGTCATACGCTGTAGTCCCACTACTGGAAAGTAAAGTGA 2804
Db 13929 AAAATTAATCCAGGATGATGTGAGCGGCTGTAGTCCAGCTACTCGGAGAGCTGAGGC 13988
QY 2805 GGTGATGTC--TTGACACAGAGTTTGAAGGCTTAAGGTGAGCATGATCACAACTGCA 2862
Db 13989 AGAGATATGCGTGAACCTGGAGGCGAGCTTGCATGAGCGAGATCACCGCACCTGA 14048
QY 2863 CTCAGCTTGAAGTGAACAGAGAAACCTGTCTCTTAAAAAGAAAGAAATGTGATTTTA 2922
Db 14049 CTCAGCTGAGCGACAGAGCGAGACTCGCTCTCAAAAAAAGTGGGTGAT 14108
QY 2923 TTCTTGAACAGTACAGTATTAATTAAGTTG 2958
Db 14109 AGGTTGCGCTTAATGATGTCAAAGTGTAAACCTG 14144

RESULT 7
PCT-US95-07201-43
; Sequence 43, Application PC/TUS9507201
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Becerra, Sofia
; APPLICANT: Patricia; Schwartz, Joan P.;
; APPLICANT: Taniwaki, Takayuki;
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07201
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/367,841
; FILING DATE: 30-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,963
; FILING DATE: 07-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22481 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: P1-147
; LOCATION:
; IDENTIFICATION METHOD:

OTHER INFORMATION: full length genomic
OTHER INFORMATION: Sequence for PDR plus flanking sequences.
PCT-US95-07201-43

Query Match 4.9%; Score 306; DB 5; Length 22481;
Best Local Similarity 55.8%; Pred. No. 3e-62;
Matches 756; Conservative 0; Mismatches 575; Indels 25; Gaps 8;

1614 TCAGAAAGTTTCAGATCGCGCGCGCGCTGCTCAGCGCTGTAATCCAGCACTTTGGG 1673
12803 TTAAGCTCAGAGGAGGCGCGCGCGCTGCTCAGCGCTGTAATCCAGCACTTTGGG 12862
1674 AGCGCGAGCGCGCTGATCAGAGCTCAGAGATCAAGAGATCCAGCTTAAACAGCTGA 1733
12863 AGGTGAGGTGGGTGATCAGAGCTCAGAGATCAAGAGATCCAGCTTAAACAGCTGA 12922
1734 AACCCGCTCTCAAAAAATATATAAATTTAGCCAGCGCTGCTGCGCGCTGTAAGT 1793
12923 AACCCGCTCTCAAAAAATATATAAATTTAGCTGGGTGTGGTGGCGCGCTGTAAGT 12982
1794 CCCAGCTACTCAGAGAGCTGAGGCGAGAGATGGCGTGAACCCGAGAGGTGAAGTTGC- 1852
12983 CCCAGCTACTCAGAGAGCTGAGGCGAGAGATGGCGTGAACCCGAGAGGTGAAGTTGA 13042
1853 GTGAGCGAGATCGCGCGCGCTGCGCTCAGCGCTGCGAGAGATGAGTGGCTTCAA 1912
13043 GTGAGCTGAGCTCGACACACTGACCTCCGCGAGAGATGAGTGGCTTCAA 13102
1913 AAAAAAAAAAAAAAAAAAGTTTCAGATCTTAAACACATGCAATTCACAGTCTAG 1972
13103 AAAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 13162
1973 AATAGAGAGAGATGTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2032
13163 TTAGAGACAGAGAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 13222
2033 TAGAGATATGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2085
13223 CTGCAATCTTGGGCTCTGAGTTCAACCAATCTCAGCTCAGCTCAGCTCAGCTCAG 13282
2086 AGGAAACACAGTGTGTTGCAATCTCCAGGTTCCATTAAGTGCATTAAGAAATCAAT 2145
13283 GACCAAGGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 13342
2146 GGTAGCAACCTGCAATTTAAATAAATAGATGAGAGAGAGAGAGAGAGAGAGAG 2205
13343 GCTTCAACATGTGTCTGAGGTGTGAGAACTCTGAGAGTCAAGATCCATCTTCTC 13402
2206 AGATGCAATTAATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2265
13403 GGCCTGC-CAGAGTGTGGAGATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 13461
2266 ACTGACTATAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 2325
13462 TATTTATATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 13521
2326 AGACTGAGTCAAAAAATCTTTAAGCACTGCGCTGCTCAATCTTATTTTGAAGAG 2385
13522 CAGGCTGAGTACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 13581
2386 GAGAAATCCAGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2445
13582 CGAATCTGTGCTCAGCTCCTGAGTACGTGGAGATTAAGGAGATCCACCAATGCCC 13641
2446 CATGCTAAAGAGATTTAGTTAAATTTCTTCCCACTAGAGGCACTGCACTTTGA 2505
13642 GTTAATTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 13694
2506 GTTCTGTGGAGCTGAGAGATCAGAGATTAATTAAGTGAAGAGAGAGAGAGAGAG 2565
13695 GGTCTGAGAGCTGAG 13751
2566 TATATTGGAGTGGTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2625

13752 -ATTATGATGAG 13810
2626 TGTGCTCAGACACTTAATATCCAGCACTTTGGAGAGAGAGAGAGAGAGAGAGAGAG 2685
13811 CGGTGCTATGCTGTATATCCAGCACTTTGGTGAAGAGAGAGAGAGAGAGAGAGAG 13870
2686 CCCAGAGTTCAAGATGAG 2744
13871 --CAGAGATTCAG 13928
2745 AAAAATCAGTGAAG 2804
13929 AAAAATCAG 13988
2805 GGTGATCGC--TTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2862
13989 AGAGAGATGGCGTGAACCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 14048
2863 CTCAGCTTGAAGTACAG 2922
14049 CTCAGCTTGAAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 14108
2923 TTCCTAGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2958
14109 AGGTTTGGCTTATAGGTCAAGAGTGTAACTTG 14144

RESULT 8
US-09-875-223-2

Sequence 2, Application US/09875223
Patent No. 6391850
GENERAL INFORMATION:
APPLICANT: No. 6391850, University of California
APPLICANT: No. 6391850, Bouck
APPLICANT: David Dawson
APPLICANT: Paul Gillis
TITLE OF INVENTION: Methods and Compositions for Inhibiting Angiogenesis
FILE REFERENCE: 0290-2303
CURRENT FILING DATE: 2001-06-06
PRIOR FILING DATE: 09/122, 079
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: PCT/US98/15228
PRIOR FILING DATE: 1998-07-23
PRIOR APPLICATION NUMBER: US 08/899, 304
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 2
LENGTH: 22484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Unsure
LOCATION: 1...22484
OTHER INFORMATION: "n" means either a, c, t, or g
US-09-875-223-2

Query Match 4.9%; Score 306; DB 3; Length 22484;
Best Local Similarity 55.8%; Pred. No. 3e-62;
Matches 756; Conservative 0; Mismatches 575; Indels 25; Gaps 8;

1614 TCAGAAAGTTTCAGATCGCGCGCGCGCTGCTCAGCGCTGTAATCCAGCACTTTGGG 1673
12803 TTAAGCTCAGAGGAGGCGCGCGCGCTGCTCAGCGCTGTAATCCAGCACTTTGGG 12862
1674 AGCGCGAGCGCGCTGATCAGAGCTCAGAGATCAAGAGATCCAGCTTAAACAGCTGA 1733
12863 AGGTGAGGTGGGTGATCAGAGCTCAGAGATCAAGAGATCCAGCTTAAACAGCTGA 12922
1734 AACCCGCTCTCAAAAAATATATAAATTTAGCCAGCGCTGCTGCGCGCTGTAAGT 1793

Db 12923 AACCCGCTCTCTAATAAATAATTAAGCTGGGTGTGGGGGGCGCTGTAGT 12982
QY 1794 CCAAGCTACTCAGAGAGCTGAGGAGAGAAATGGCGTGAACCCGGGAGGTAGATTGC- 1852
Db 12983 CCCAGCTACTCGGAGAGCTGAGGAGAGAAATGGCGTGAACCCGGGAGGTAGATTGCA 13042
QY 1853 GTGAGCCGAGATGCGCCGCTGAGGAGAGAAATGGTTTTCAGAAAGTACAGATGCTGCA 1912
Db 13043 GTGAGCTGAGCTGACACTGACCTCCAGCTGGGGGAGAGATGAGACTCGCTTCAA 13102
QY 1913 AAAAAAAAAAAAAAAAAAGTTTCAGATCTTAAACATGATTTTCAACAGTCTAG 1972
Db 13103 AAAAAAAAAAAAAAAAAAGAAAGAAAGAAAGAAAGTCTTTTCTTTTAACTTTCTTTT 13162
QY 1973 AATAAGAGAGCATGTTACAGAGAGAGAAATGTTTTCAGAAAGTACAGATGAGGAAA 2032
Db 13163 TTGAGAGCAGAGCTCACTCCATCAACCAGCTGAGATACAGATGAGTGCATCTTGGCTCA 13222
QY 2033 TAGAGATATGTTCAAG 2085
Db 13223 CTGCAATCTTGGCTCTCTGAGTTCAACCAATTTCTCATGCTCCAGCTCCCAATAGCTGG 13282
QY 2086 AGGAAACACAGTGTGTTGCAATCTCCAGTTCCATTAGTCCGTTAGAAATCAATATGT 2145
Db 13283 GACCAAG 13342
QY 2146 GGTTAGCACTGATTTTAAAAATGAATTAATGATGAGAGAGAGAGAGAGAGAGATAT 2205
Db 13343 GCCTCAACAGTGTGCTCAGGTGTGTGAACTCTGAGCTCAAGATGATCATCTTCTC 13402
QY 2206 AGCATGATTAATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2265
Db 13403 GGCCTGC-CAAGTGTGAGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 13461
QY 2266 ACTGATTAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2325
Db 13462 TATTTATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 13521
QY 2326 AGAGTCAAGTCAAAATCTTTTAAAGACATGCGCTGTCTTAATCTTTTATTTTGAAG 2385
Db 13522 CAGGCTGAGATGAG 13581
QY 2386 GAGAAATCCAAAGCTGAG 2445
Db 13582 CGAATCTCGCTCCTGAGCTCCTGAGTGTGAGATTAAGAGAGAGAGAGAGAGAGAG 13641
QY 2446 CATTTGCTAAAGATTTCTTAAATTTCTTCCACTAGTGCATATGCACTTCTA 2505
Db 13642 GTTAATTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 13694
QY 2506 GTTCTGTGGCTGGAATACAGAAATATTTATTTAGGAAACAGATACAGATCTGGGAAA 2565
Db 13695 GGTCTGAACTCTGACCTCAAGTGTATCAACCACTCAGCTCCCAAAATGTGTGG--- 13751
QY 2566 TATATTTGGTGTGCTGAG 2625
Db 13752 -ATTATTAAGTGTAGGCACTCGGCTGATGTTTAAAAAGGTGTGTATGGGGCTGGCG 13810
QY 2626 TGGTGTCTACACTTAATTTCAAGCATTTGGAGAGCCAAATCAAGAAATGCTTGA 2685
Db 13811 CGGTGTCTATGCTGTAAATCCAGACATTTGGTAAACGAGGCGGGTGTATCAAGGT 13870
QY 2686 CCAAGAGTTCAAGATGAG 2744
Db 13871 --CAGAGATCGAGACATCTCTCAACAGGTGAACCCGCTCTCTAATAAATACA 13928
QY 2745 AAAAAATCAGCTGAG 2804
Db 13929 AAAAAATTAACAGAGATGAG 13988
QY 2805 GGTGATTCG--TTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2862
Db 13989 AGGAGATGAG 14048

QY 2863 CTCACCTTGAGTACAG 2922
Db 14049 CTCACCTTGAGTACAG 14108
QY 2923 TTCCTTAGCAGTACAGTATGATGATGATGATGATGATGATGATGATGATGATG 2958
Db 14109 AGGTTTGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 14144

RESULT 9
US-09-875-114-2
; Sequence 2, Application US/09875114
; Patent No. 6670333
; GENERAL INFORMATION:
; APPLICANT: No. 6670333chwestern University
; APPLICANT: No. 66703331 Bouck
; APPLICANT: David Dawson
; APPLICANT: Paul Gilis
; TITLE OF INVENTION: Methods and Compositions for Inhibiting Angiogenesis
; FILE REFERENCE: 0290-2302
; CURRENT APPLICATION NUMBER: US/09/875,114
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 09/122,079
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: PCT/US98/15228
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: US 08/899,304
; PRIOR FILING DATE: 1997-07-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 22484
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: 1...22484
; OTHER INFORMATION: "n" means either a, c, t, or g
US-09-875-114-2

Query Match 4.9%; Score 306; DB 4; Length 22484;
Best Local Similarity 55.8%; Pred. No. 3e-62;
Matches 756; Conservative 0; Mismatches 575; Indels 25; Gaps 8;

QY 1614 TCAAGAGTTTCAGATTCGGCGCGCGGCTGCTCAACGCTGTAAATCCAGACATTTGGG 1673
Db 12803 TTAAGACTGACAGTGAAGGCGCGGCGGCTGCTCAACGCTGTAAATCCAGACATTTGGG 12862
QY 1674 AGCCGAGCGCGGTGATCAACAAGGTCAAGAGATCAACATCTGTTTAAACAGGTGA 1733
Db 12863 AGGTGAGGTGTGTATCAACAAGGTCAAGAGATCAACATCTGTTTAAACAGGTGA 12922
QY 1734 AACCCGCTCTCAAAAAAATTAATAAATTAAGCAGGCGGTGTGCGCGCTGTAGT 1793
Db 12923 AACCCGCTCTCAAAAAAATTAATAAATTAAGCAGGCGGTGTGCGCGCTGTAGT 12982
QY 1794 CCAAGCTACTCAGAGAGCTGAGGAGAGAGAGATGCGGTGAACCCGGAGAGTGAATTGC- 1852
Db 12983 CCAAGCTACTCAGAGAGCTGAGGAGAGAGAGATGCGGTGAACCCGGAGAGTGAATTGC 13042
QY 1853 GTGAGCCGAGATGCGCCGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1912
Db 13043 GTGAGCTGAGCTGACACTGACCTCCAGCTGGGGGAGAGATGAGACTCGCTTCAA 13102
QY 1913 AAAAAAAAAAAAAAAAAAGTTTCAGATCTTAAACATGATTTTCAACAGTCTAG 1972
Db 13103 AAAAAAAAAAAAAAAAAAGAAAGAAAGAAAGAAAGTCTTTTCTTTTAACTTTCTTTT 13162
QY 1973 AATAAGAGAGCATGTTACAG 2032
Db 13163 TTGAGAGCAGAGCTCACTCCATCAACCAGCTGAGATACAGATGAGTGCATCTTGGCTCA 13222

QY	2033	TAAAGATATGTTCAAGGAAGAGAGACCCCAAGATCATGTTGT-----TAGGGTTAG	2085
Db	13223	CTGCATCTTGGGCTCTGTAGTTCAACCAATCTCATGCTCAGCCTCCCAATAGCTGG	13288
QY	2086	AGGAAACACAGTGTGTCATCTCCAGTTCATTAGTCGTTATGAATCAATATGTG	2145
Db	13283	GACCAACAGGACGATGCCACCAAGCCAGCTAATTTTGGGGTATTTTATAGTAGATGG	13342
QY	2146	GGTTAGCACTGCTATTTAAAAATGAAATTAATGATGAGAAAGAAATAGAAATAT	2205
Db	13343	GCCTCACCAAGTGGCTCAGGTTGGTGTGAAACTGTAGCTCAAGATCATCTTCC	13402
QY	2206	AGCATGATATACATTTTGAAGAAGCAAGATATATTTTCTGCACTTTTGCTCCAAATGTA	2265
Db	13403	GGCTGTC-CAAAGTCTGGGATTAATAGCATTAAGCCATGACCTTACCTGCCAATTTT	13461
QY	2266	ACTGTACTATATTTTATGTATGATGTGAATACAGATACATATATATTTCTTACTGT	2325
Db	13462	TATTTATATTTATTTTATTTATTTATTTATTTTGAACAGGGCTCAGCTGTGCAC	13522
QY	2326	AGACTGCACTCAAAAAATCTTTAAAGCATGCGCTGTCTAATCTCTTATTTTGACAG	2385
Db	13522	CAGCTGAGATACAGTGGCACTATCTCAGCTCATCTCAACCTTGCCTCTGGGTTCAAG	13581
QY	2386	GAGAAATCCAAATCTGAGAGACAACAATTTGGCGTAGGTATAGAAACAGCTTATGC	2445
Db	13582	CGAATCTCGTCCCTCAGCCTCTGATGATGCTGGGATTAACGCAATGACACCAATCC	13641
QY	2446	CATTGCTAAAAAGTATCTTAGTTAAATCTTTCCCATAGTGCATCTGCACTTCTTA	2505
Db	13642	GTTAATTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	13699
QY	2506	GTTCTGTGGCTCGAAATACAGAAATATTATTAGTGAACAGCATACCAAGTCTGGGAAA	2565
Db	13695	GGCTCGAACTCTGTGACCTCAAGATATTCACCACTCAGCCTCCCAAGTGTGG---	13753
QY	2566	TATATTGGATAGGAGGCTGAGAGCTCATTTTCTAAGAAATGAGACTTAGGCAAGTGA	2625
Db	13752	ATTATATAGTGTAGACCACTGGCTGATGTTTATTTAAAGTGGATCATGGGCTGGGCG	13811
QY	2626	TGGTGGCTCACACTTATTAATTCAGACACTTTGGAGGCGCAAGTCAAGAAATGCTTGA	2685
Db	13811	CGGTGGCTCATGCTGTGAATCCAGACACTTTGGTGAACGAGGCGGGTGGATACAAAGT	13870
QY	2686	CCCAAGATTCAAAGTACAGATAGCATGGGCAACATACAGAACTCATCTCTCAAAAAAT-TT	2744
Db	13871	--CAGGAGATCGAAGACCATCTCTCTPAACAAGTGAACCCCGTCTCTAATAAATACA	13928
QY	2745	AAAAATCAGTGAAGCATGTGGCATACGCTGTATGTCCTCACTAAGGAGAGTATGAGT	2804
Db	13929	AAAAATTAACCAAGGCAATGATGGGGGCTGTATGTCACAGCTATCTGGGAGGCTAGGC	13988
QY	2805	GGTGAATCGC--TTGACACAGAGATTGAGGCTAAAGTGAAGCATGATACAACTGCA	2862
Db	13989	AGGAGAAATGGCTGTAACTTGGGAGGCGGAGCTTGCACTAGCCGAATCAACGACCTGA	14044
QY	2863	CTCCAGTTGATGACAGAGAAAGACCTGTCTCTAATAAAGAAAGAAATGTGATTTTA	2922
Db	14049	CTCCAGCTAGCGACAGACGAGACCTCGCTCTCAAAAAAATAAATAAAGTGGTCAAT	14108
QY	2923	TTCCTTAGACATACATCTATTATCTAATTAAGTTTG	2958
Db	14109	AGGTTTCGGCTTATAGGTCAAGATGTTTAAACCTG	14144

RESULT 10
US-08-781-891-79
Sequence 79, Application US/08781891
Patent No. 6090620
GENERAL INFORMATION:
APPLICANT: Fu, Yang-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Ohnima, Junko

APPLICANT: Mulligan, John T.
 APPLICANT: Schellengberg, Gerald D.
 TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
 TITLE OF INVENTION: WERNER'S SYNDROME
 NUMBER OF SEQUENCES: 209
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED and BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104-7092
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/781,891
 FILING DATE: 27-DEC-1996
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Mr. 6090620cemburg Ph.D., Carol
 REGISTRATION NUMBER: 39,317
 REFERENCE/DOCKET NUMBER: 240052.419
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 79:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 87350 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

Query Match	Score	DB	Length
4.9%	305	3	87350

Best Local Similarity 52.6%; Pred.NO. 9.9e-62;
Matches 777; Conservative 0; Mismatches 695; Indels 5; Gaps 5;

QY	458	TTTGATACATTTGGAAATATATATTAAGTTGTGAAGCTGGCGGGTGGCTCAAGCCCTGTATCCCA	517
Db	8934	TGGATATTTTCTTAAAACTATGTATTGGGCGGGGTGCAGTGGCTCATGCTCTGTATCCCA	8993
QY	518	GCCCTTTGGAGGCGCAGATGGCGCAGATCATGAGGTGAGGATTGTGAGCCAGCCTAGCC	577
Db	8994	GCACTTTGGGAGGCGCAGCGGGGTGATCCGAGGTGAGGAGATCGAACCATCTGGACT	9053
QY	578	AACATGCTGMAACCCCGCTCTTACTATAAGTATCAAAAAATTAGCCAGTGTGTGGCCCA	637
Db	9054	AACATGATGMAACCCCGCTCTTACTATAAAATACAAAAAATTAGCCGAGTGTGTGGCCAG	9113
QY	638	CACCTGTAAACCCGAGCTGTGAGAGAGATGGCAGGAGAAATTCCTTGAAACCAAGAGGC	697
Db	9114	TGCTGTATGTCCGAGTACTCGGGAGGCTGAGCGAGGAATGTGTGMAACCAAGAGGC	9173
QY	698	AGAGGTCCAGCGAGCAAAAGTACACCAATGCACTGTATAGCTTGGATGACAGGGCAACAT	757
Db	9174	GGATTTGCAAGTGAAGCCGAGATCGTGCCATGCACTCCAGCCTGGGCAACAGGAGAGCT	9233
QY	758	CCAACCTCAAAAAAAAAAAAAAAAAAAAAAATAATGTATGTGTGTCTATTAACAATTAATAGG	817
Db	9234	CTGTCTCAAAAAAAAAACCAACCACTGAACAACAAAAAACTAATAAACAABAAAACAABA	9293
QY	818	CAGTGAAGAGCAAGTGTCTAAAGCTTGTGACATGTGTAACTAGGAATATCTGTGGAAACAC	877
Db	9294	AACATATATATTAGCATGGGTGTGGCAAACTATAGCCTGTAGGCAAAATCTGCATGCTGTT	9353
QY	878	ATATATAAGGAACCTTAACCCAGTCTCTGGAAGTAAAGTTTGGAAAGAAATGTTTGGAGAC	937
Db	9354	TTATTTTTTTTTTTTTTTTGATCATAGGGTCACTACAGGCT-GTCACACAGGCTGAGAGGC	9412
QY	938	AAAGGTTTAAAGAGAGTGAIAAAAAAAAAATTAAATACCACTTTAGCTGTGTGAGAAATGG	997

Db	9413	AGTGGATATCATAGCTCAGTGAACCTCAAAATTCCTGGGCTCAGCAATCTCTTCC	9472
QY	998	GATTAGGAGCTAACTAGAGAAATCAAATAGGAATCTTTCATGTATGTTAAGAACCCTGG	1057
Db	9473	TCACCTCAGCTTCCCAAGTAGCTACAGGCATGCACTACCAACCCAGTTAATTTAAACAA	9532
QY	1058	TAAGGATGAAGACCAATTCATTTATCTGACCAATCGCGGACTTTTTTTTATGTAAATGC	1117
Db	9533	ATTTTTTTTTGGTAAAGACAGCTCAGTATGTGTCCAGCGCTGGTTTTTCAACTCTTGC	9592
QY	1118	TTGGCAATTTAATATAGAGACAGAAATGTAACAATGGATTGAGTCAGATGTAAGT	1177
Db	9593	CTCAATCAGTCCCTCTACTTCCAGCTCCTTAAGTGTGGATTAATAGGCTCGAGCATCA	9652
QY	1178	TCCTGCAGACATGTGAAGAAAGACAGATAGCAAGATGGTAAGATTTACAGA	1237
Db	9653	CGCTTGACTAATGTTTTTTTGTAAATBAATGTTTTCTCAGAACACAGCCATGCTTTTGTTTA	9712
QY	1238	CAGAAATTAAATGTCTGCGCCAGTGGCATCTAGCTGAGTCTAATCTGAGGAAAGAAATG	1297
Db	9713	TGTGTTATGTAGGGCT -GCTGAGTTAATGTTAGTTGGCTACAAACCTATCATGCGCTATA	9771
QY	1298	AAGATTAACAGCTTCTGATAGTTATGAAGAGATGGAAGCTTCAAGACCTTACAGATG	1357
Db	9772	AAGCTGAAATACCTTACTATCTGTGCTCTTTTAAAGAAATGTTTTCTACCTCTGACTAGA	9831
QY	1358	TTGATTAATATGAAGAATGATTGGAAGAAAGATTAATCTGTAGAGAGTGAATTTTCAGGC	1417
Db	9832	CTAGCTGTCTCAAAATCTTCAATGAATTTGGAAGTTTTCTACCACTTTTCTGACCA	9891
QY	1418	TTGAGTCACTCTCACATACCAACACTGTGCTAAATGCTTCAAGACATGATCCCTGACC	1477
Db	9892	TAATGCACTTAGTTAGAAATGAATTAAGACATTAACAACAAATCTCATGACTTTTGA	9951
QY	1478	TCAAGGGACTTACAGCCAAACAAAGATGAAGAAATACACCAATACTATATAGAC	1537
Db	9952	AATTAATAATTAACCTTAATTAATTCATATTCAAAGAAATAATTAACCTGAATTTAA	10011
QY	1538	ACTGTGTGAATATCAAGAAAGAAATACGATCTAGTACTGTAGATGTGCAAGCGCATCA	1597
Db	10012	AAAAATTTTAACTTACAGATTAACATACTAATTAATGATTAACATTTTAAACATTGAGGA	10077
QY	1598	AAGATATCTTCTAGTTTCAGAAGTTTCAGATCGGCGGGCGCGGTGCTCAGCCTGTGA	1657
Db	10072	TAGTTACATATATATACATTAAACCTGATAGAGAGCGGTGGTGGCTCAGCGCTGTGA	10133
QY	1658	ATCCAGCACTTTGGAGGCGGAGCGGGGTGATCAACAAGTCAGAGATCAAGACCATC	1717
Db	10132	ATCCAGCACTTTGGAGGCGGAGCGGTGGGTGATCAAGAGTCAAGAGATTGAACCATC	10199
QY	1718	CTGGTTAACACGCGTGAACCCCGTCTTACAAAATAATATAAAATTTAGCCAGCGGTGT	1777
Db	10192	CTGGCCAAACATGGTGAATCTCCGCTCTCTACTATAAAATA - CAAAATCAGCTGGGCGGTGT	10255
QY	1778	GGCGGGCGCCCTGTAGTCCAGCACTACAGAGGCTGAGAGAGAGAAATGGCCGTGAACCCG	1837
Db	10251	GGCAACGCGCCCTGTAGTCCAGCACTTGGAGGCTGAGAGAGAGAAATGCTTGAACCTGG	10310
QY	1838	GGAGGTAGAAGTTTG -CGTAGCCGAGATTCGCGCACTGCGCTCAGCTTGGGCGGACAGAG	1896
Db	10311	GGAGGCGGAGGTTGCGGTGAGCCGAGATTTGGGCGCATGCACTCAGCGCT -GGGACAGAG	10366
QY	1897	TGAGACTGCGCTCAAAAAAAAAAAAAAAAAAAAAA 1933	
Db	10370	CGACACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAA 10406	

RESULT 11
US-09-618-166-79
; Sequence 79, Application US/09618166
; Patent No. 6583112
; GENERAL INFORMATION:

	APPLICANT: Fu, Ying-Hui	
	Yu, Chang-Fn	
	Oshima, Junko	
	Mulligan, John T.	
	Schellenberg, Gerald D.	
	TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO WERNER'S SYNDROME	
	NUMBER OF SEQUENCES: 209	
	CORRESPONDENCE ADDRESS:	
	ADDRESSEE: Seed Intellectual Property Law Group	
	STREET: 701 Fifth Avenue, Suite 6300	
	CITY: Seattle	
	STATE: Washington	
	COUNTRY: USA	
	ZIP: 98104-7092	
	COMPUTER READABLE FORM:	
	MEDIUM TYPE: Floppy disk	
	COMPUTER: IBM PC compatible	
	OPERATING SYSTEM: PC-DOS/MS-DOS	
	SOFTWARE: Patentin Release #1.0, Version #1.30	
	CURRENT APPLICATION DATA:	
	APPLICATION NUMBER: US/09/618,166	
	FILING DATE: 17-Jul-2000	
	CLASSIFICATION: <Unknown>	
	ATTORNEY/AGENT INFORMATION:	
	NAME: Mcmasters, David D.	
	REGISTRATION NUMBER: 33,963	
	REFERENCE/DOCKET NUMBER: 240052.419C1	
	TELECOMMUNICATION INFORMATION:	
	TELEPHONE: (206) 622-4900	
	TELEFAX: (206) 682-6031	
	INFORMATION FOR SEQ ID NO: 79:	
	SEQUENCE CHARACTERISTICS:	
	LENGTH: 87350 base pairs	
	TYPE: nucleic acid	
	STRANDEDNESS: single	
	TOPOLOGY: linear	
	SEQUENCE DESCRIPTION: SEQ ID NO: 79:	
	US-09-618-166-79	
Query Match	4.9%; Score 305; DB 4; Length 87350;	
Best Local Similarity	52.6%; Pred. No. 9.9e-62;	
Matches 777; Conservative	0; Mismatches 695; Indels 5; Gaps 5;	
QY	458 TTGTACATTTTGGATATATATAAGTTGTAAGCTTGCGCGGTGGCTCACGCCTGTATATCCCA	517
DB	8934 TGGAATATTTCTAAAACTATGTATTTGGCGCGGTGTCAGTGGCTCATGCCTGTATATCCCA	89933
QY	518 GCCCTTTGGGAGGCCACAGATGGGCGAGATCATGAGGTGAGGATTGAGCACAGCTTAGACC	577
DB	8994 GCATTTTGGAGGCGCGAGCGGGTGGATCCCGAGGTAGAGATCGAACCATCTTGACT	9053
QY	578 AACATGTGAAACCCGCTCTCTAATAAGATACAAAATAATGACAGTGTGTGGCGCA	637
DB	9054 AACATGTATGAAACGCCGCTCTCTAATAAATACAAAATAATGACCGATGTGGTGGAG	9113
QY	638 CACCTGTAAACCCAGCTGCTCGAGAGAGTGAAGCAGAGAATGCTTGAACCCAGAGGCG	697
DB	9114 TGCTGTATAGTCCACGCTACTCTGGGAGGCTGAGGAGGAGATGATGTGAACCCAGAGGCG	9173
QY	698 AGAGTGCACGACGACAAGATGCAACCAATGCACTGTAGCCTGTGATGACAGGCGCAAGACT	757
DB	9174 GGAGTTGCAAGTGAAGCGAGATCGGCCACTGCCAGCTTGGGCAACGAGCGAGACT	9233
QY	758 CCAACTGAAAAAAAAAAAAAAAAAAGAAATATGTAACTGTGCTATPACAAATPAATAGG	817
DB	9234 CTGTCTGAAAAAAAAAACCAACCACTGAAACAAACAAAAAACTAAAAAAACAAAAACAAAA	9293
QY	818 CAGTGAAGACAAAGTCTTAAAGCTATGACCAATGTAAGTAAGTAAGTAAGTAAGTAAGTAAG	877
DB	9294 AACTATATATTTAAGCATGGGTGGCAAACTATGCTGTAGGCAATCTGCATGCTGTT	9353
QY	878 ATATAAAGGAGAACCTAACCGAGTCTGTGAAGTAAGTTTTGGAAGAAATGTTTGAAGAC	937

Db	9354	TTATTTTTTTTTTTTTTTTGACATAGGTCACATACAGGCT-GTCACACACAGCTGGAGAC	9412
Qy	938	AAAGGTTTAAAGAGTGAACAAAAAATTTAAATACAGTTTACCTGTGGAGATGG	997
Db	9413	AGTGGTATGATCATAGCTCACCTGAACCTCAAAATCTCGGGCTCAAGCAATCTCTTGGC	9472
Qy	998	GATAGGGAGCTAATAGAGAAATCAATTAGGAATGTTTCAATGGATAGTTAAGAACCTCG	1057
Db	9473	TCACCTCAGCTTCCCAAGTAGCTACACAGCATGCACTACACAGACCAGTTAATTTAAACAA	9532
Qy	1058	TAAAGGTGAAGACCATTAACATTTATCTCACCATCGCGGAGCTTTTTTTTATAGTAAATGC	1117
Db	9533	ATTTTTTTTGGTAGAAGACAGCTCTCAGTATGTTGCCAGGCTGGTTTCAAACTCCTTGC	9592
Qy	1118	TTGGCAATTTAAATAGAGAGACAGAAATGTAGACAGTTGGAATTTAGCTCAGAGTTGAAGT	1177
Db	9593	CTCAATCAGTCTCTCTACTTCAAGCCTCTTAAAGTGTGGATTTATAGGCTTGAGCCATCA	9652
Qy	1178	TCATCCAGACATGTGAAGAGAGACAGGTAGGCACAGAGATTTGAAGATTTACAAG	1237
Db	9653	CGCTTGACTAATGTTTGTGAAATAAAGTTTCTCAGAAACAGCCATGCTTTTGTTTA	9712
Qy	1238	CAGAAATTAATGTCGTGGCAGTGGCATCTAGTCTGAGTCTAATCTGAGGAGAAAGT	1297
Db	9713	TGTGTTATGTAAGGCT-GCCTGACTTAAGTAGTTGGCTTACAAAGCCTATATAGCCTATA	9771
Qy	1298	AAAGTAAAGCACTTGTGATAGTTATGAGAGAGTGAAGGCTTCAAGCATTAAGATG	1357
Db	9772	AAAGCTGAAATACTTACTATCTGTGCTTTATAGAAAGTGTTTTCTGACCTGTACTAGA	9831
Qy	1358	TTGATTTAAATGAAGAAATGATTTGAGAAAGAAATACCTGTAGAGAGTGAATTTCAAGC	1417
Db	9832	CTAGCTGTCTCAAAATTTCTTCATGAAATTTGGAAAGTTTCTCAGCAATTTTCTACCA	9891
Qy	1418	TTGAGTACTCTCACATACAGACACATGTCCTAAATGTCCTTCAAGACATGATCCCTGGCC	1477
Db	9892	TATGACTTGAAGTTAAGATTAATTAAGACAGATTAACACAAATCTCATGCAATTTGGA	9951
Qy	1478	TCAAGGACTTACAGCCAAAAACAAGATTAAGAAATACACCAATACTATTAATAGAC	1537
Db	9952	AATTAATAATACCTTAATTAATTCATATTTCAAGAAAAATCAAACTGGAATTTAAA	10011
Qy	1538	ACTTGTGATATATCAAGAAAGAAATACATCTAGTACTGTAGATGTGCAACGGCATCA	1597
Db	10012	AAATTTTAAACCTACAGATTAACCTAATTAATGACTTAAACATTTTAAACCTTAGGG	10071
Qy	1598	AAGATATCTTCTAGTTTCAGAAAGTTCAATCGCGCGCGCGGTGGCTCAGCCCTGTA	1657
Db	10072	TAGTTAACATGATATTAATTAATACTGTAAAGGCTGGGTGGGTGCTCAGCCCTGTA	10131
Qy	1658	ATCCACAGACTTTGGAGGCGAGGCGGATGATCACAAGGTCAAGAGATCAAGACATC	1717
Db	10132	ATCCACAGACTTTGGAGGCGGAGGCTGGGGGATCACAGGTCAAGAGATTTGAACATC	10191
Qy	1718	CTGTGTTAACACGAGTGAACCCCGTCTCTACAAAAATATATAAAATTTAGCCAGCGTGGT	1777
Db	10192	CTGSCAAACATGGGAATCCCGTCTCTACTATAAAAAA-CAAAATCAGCTGGGCGTGGT	10251
Qy	1778	GGCGGGGCGCTGTAGTCCCACTACTCAAGAGGCTGAGGCGAGAGAAATGGCGTGAACCG	1837
Db	10251	GGCAGCGGCGCTGTAGTCCCACTACTTGGGAGGCTGAGGAGAGAAATCCCTTGAACCTG	10311
Qy	1838	GGAGGTAGAGTTT-CGTGAAGCGAGATCGCGCACTGCGCTCAAGCTGGGCGACAGAG	1896
Db	10311	GGAGGCGAGGTTCCCTGAGCCGAGAGTATGGGCGCATCTGCATCGAGCT-GGCGACAGAG	10361
Qy	1897	TGAGACTGCGTCTCAAAAAAAAAAAAAAAAAAAAAA	1933
Db	10370	CGACACTCTTGTCTCAAAAAAAAAAAAAA	10406

```

US-08-520-373D-4
; Sequence 4: Application US/08520373D
; Patent No. 6451763;
; GENERAL INFORMATION:
; APPLICANT: Tombran-Tink, Joyce
; APPLICANT: Steele, Fintan R
; APPLICANT: Chader, Gerald J
; APPLICANT: Beceerra, Sofia P
; APPLICANT: Johnson, Lincoln V
; APPLICANT: Rodriguez, Ignacio R
; TITLE OF INVENTION: RETINAL PIGMENTED EPITHELIUM DERIVED NEUROTROPIC FACTOR
; FILE REFERENCE: 2026-4203U51
; CURRENT APPLICATION NUMBER: US/08/520,373D
; CURRENT FILING DATE: 1995-08-29
; PRIOR APPLICATION NUMBER: 08/377,710
; PRIOR FILING DATE: 1995-01-25
; PRIOR APPLICATION NUMBER: 08/279,979
; PRIOR FILING DATE: 1994-07-25
; PRIOR APPLICATION NUMBER: 07/894,215
; PRIOR FILING DATE: 1992-06-04
; PRIOR APPLICATION NUMBER: 07/952,796
; PRIOR FILING DATE: 1992-09-24
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 14581
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: mRNA: 6683; EXON: 6683-6790; EXON 11584-11675;
; OTHER INFORMATION: EXON: 14539-14581; INTRON: 6791-11583; INTRON:
; OTHER INFORMATION: 11676-14538; CDS: 11584-11675; 14539-14580
US-08-520-373D-4

Query Match 4.9%; Score 304.4; DB 4; Length 14581;
Best Local Similarity 55.7%; Pred. No. 5,8e-62;
Matches 755; Conservative 0; Mismatches 576; Indels 25; Gaps 8;

QY 1614 TCAAGAA GTTTCAGANTCGCGCGGCGGTGCTCAGCGCTGTAATCCAGACCTTTGGG 1673
DB 12814 TTAAGACTGCAAGTAGAGGGCGCGGGCGGTGCTCAGCGCTGTAATCCAGACCTTTGGG 12872

QY 1674 AGCGCCGAGGGGGGTGATCTCAAGGTTCAGAGATCAAGACCAATCCGTGTAACACGGTGA 1733
DB 12874 AGGTGAGGGGGGTGATCTCAAGGTTCAGAGATCAAGACCAATCCGTGTAACACGGTGA 12933

QY 1734 AACCCCGTCTCTCAAAAAAATTAAAAAAATTAGCCAGCGGTGGCGGGCGCTCTAGT 1793
DB 12934 AACCCCGTCTCTCAAAAAAATTAAAAAATTAGCTGGGTGTGTGGGGGGCGCTCTAGT 12999

QY 1794 CCCAGCTACTCAGAGAGGCTGAGGCAGAGAGAAATGGCGTGAACCCGGAGGTTAGTTGC- 1852
DB 12994 CCCAGCTACTCAGAGAGGCTGAGGCAGAGAGAAATGGCGTGAACCCGGAGGTTAGTTGCA 13055

QY 1853 GTGAGCCGAATGCGCCCACTGCGCTTCAGCCTGGGGCGACAGAGTGAAGTGGCTTCAA 1912
DB 13054 GTGAGCTGAAGCTGCGCACCACTGCACTTCAGCCTGGGGCGACAGAGTGAAGTCCGTCTCAA 13113

QY 1913 AAAAAAAAAAAAAAAAAAAAAAAAAAATTTTCAAGATCTTAAACAACACTGCATTTTCAACAGTCTAG 1972
DB 13114 AAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTGTCTTTT 13177

QY 1973 AATAGAGACATGTTATCAGAGGAGAGAAATGTTTTTCAAGAAAGTATCAGAGTAGGAGAA 2032
DB 13174 TTAAGAGACAGAGTCTCACTCCATCAACCAAGTGAAGTACAGTGGTGCATCTTTGGCTCA 13233

QY 2033 TAGAGATATGTTCAAGAAAGAGAACCCCAAGACTCATGGTTTCT-----TAGGGTTAG 2085
DB 13234 CTGCAATCTTGGGCTCTCTGAGTTCAACCAATTTCTCATGCTCAGCTCCCAATTAAGCTGG 13299

QY 2086 AGAGAAACACAGTGTTTTTCAGAAATTCACAGGTTTCAATTAGTGCATTAGAAATCAATTAGGT 2145
DB 13294 GACCAACAGCAGCTGCGCACCAAGCCCAAGCTAATTTTGGGTATTTTGTAGTAGATGGG 13355

```

QY 2146 GGTAGCAACCTGCATTTTAAAAAATGAATTAATGATGAGAGAAATGAAAAATATT 2205
DB 13354 GCCTCAACCATGCTCAGGTTGGTCTGAAACCTCGAGCTCAAGTATCATCTTCCCTC 13413
QY 2206 AGCATCATTAACATTTTGAAGAGCAAGTATTTTTCGCACTTTTGTCCAAATTGA 2265
DB 13414 GGCCTGC-CAGAGCTGGGATTAATAGGCAATAGCCATGACCTGACCTCCCAATTTTA 13472
QY 2266 ACTGTAATTATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2325
DB 13473 TATTTATATTTATTTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTT 13532
QY 2326 AGATGCAATGCAAAAAATCTTTAAGCACTGCGCTGCTAATCTTCTTATTTTGAAGAG 2385
DB 13533 CAGGCTGAGTACAGTGCATCTCTCAGCTCAACCTTCGCTCCCTGGGTTCAAG 13592
QY 2386 GAGAAATCCAGATCTGAGAGCAAAACATTTTGGCGAGTTATAGAACCACTTATGC 2445
DB 13593 CGAATCTCGCTCAGCTCCTGAGTACGAGGATTAAGGATCACCACCATGCCCC 13652
QY 2446 CATGCTAAAGATGATCTTATGATTAATTTCCCACTAGTCCATATCTGCATTTCTA 2505
DB 13653 GTTAATTTTGTATTTTATTTATTTAGTGAAGAGGGTTTC-----ACGTTGTCAGAGAT 13705
QY 2506 GTTCTGTGGCTGAAATACAAATATTTATGTAAGACAGCATACAAAGTCTGGGAAA 2565
DB 13706 GGTCTGAACTCTGACCTCAAGTGAATCCCACTCAGCTCCCAAGTCTGGG--- 13762
QY 2566 TATATGGGATGAGTGGCTGAGAGCTCATTTCTAAGAAATGAGACTTATGGCAGAGTA 2625
DB 13763 -ATTATAGGTGATGAGCACTCGGCTATGTTTAAAGATGAGGATGAGGCTGGGCG 13821
QY 2626 TGGTGGCTCAACCTTAATTTCCAGCACTTTGGAGGCGCAAGTCAAGAAATGCTTGA 2685
DB 13822 CGTGGCTATGCTCTTAATCCAGCACTTTGGTGAACCGAGGCGGATGATCAAGGT 13881
QY 2686 CCCAAGATTCAGATTAACATGAGGAGCAATAGCAAGACTTCTCTCAAAAAAT-TT 2744
DB 13882 --CAGAGATTCAGACATCTGCTTAACAAGGTGAACCCCGTCTCTCAAAAAATACA 13939
QY 2745 AAAAATCAGCTGAGCTGAGTGGCATGCGCTGATGCCACCTACTCTGGGAAGCTAGTG 2804
DB 13940 AAAAATTAACCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 13999
QY 2805 GGTGATGCG--TTGACACAGAGTTTGAAGCTTAAGTGAAGCCATGATCAACAACCTGA 2862
DB 14000 AGAGAGATGCGGTGAACCTGGGAGGCGGAGGCTTGCAGTGAAGCCAGATCAACGCACTGA 14059
QY 2863 CTCAGCTTGAAGTGAAGAGAGAAACCTGCTCTTAAAAAAGAAATGATGATTTTA 2922
DB 14060 CTCAGCTTGAAGTGAAG 14119
QY 2923 TTCCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2958
DB 14120 AGGTTTCGGCTTATAGTCAACAAGTGTTTTAAACCTG 14155

RESULT 13
US-09-544-398B-8
; Sequence 8, Application US/09544398B
; Patent No. 6770461
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Recker, Robert R.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/09/544,398B
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13

QY 486 GCTGGCGGCTGCTCAAGCTGTAATCCAGCCCTTTGGAGAGCCAGAGTGGCAGAT 545
DB 20381 GCGGGGTGGGTGGCTCAGGCTGTAATCCAGCACTTGGAGCCAGAGGAGATCC 20440
QY 546 CATGAGCTCAGAGATTTTGAAGACCACTAGCCATATGATGAACCCGCTCTCTACTAA 605
DB 20441 CTTGAGCTTGAAGATTTGAGACCACTGAGCAAT-GTGAAGCCCGTCTCTACAAA 20499
QY 606 GATCAAAAAATTTAGCCAGGTGTTGGGCGCACCTGTAACCCAGCTGCTCGAGAGAG 665
DB 20500 AATAC-AATTAATTCAGAGTGTGTGGCGCACTCTGATGCCAGCTACTCGGAGGC 20558
QY 666 TGAGGACAGAGATTTGTTGAACCCAGAGGCGAGG-TGACGAGCAAGATCAACAC 724
DB 20559 TGAGGCGGAGATGCTTGAAGCTTGAAGGAGGAGGAGGCTGAGTATGATGAC 20618
QY 725 AATGACATGATGCTGATGAGAGGCAAGCTCAATTCAAAAA-----AAAAA 782
DB 20619 AATGACATGACCTGAGGAGGAGAGAGTGAAGTCTGCTCAAAAAAAGTTAAAAA 20678
QY 783 GAAATATGATGATGCTGATTAACAATTAATAGGCACTGAGAAACAAAGTCTAAAGC 842
DB 20679 AAAAAATCTGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 20738
QY 843 TATGACATGCTACTGAGATATCTGTGGGAAACATATTAAGGAACTTAACCACTCC 902
DB 20739 GTGCCAATTTGCTATTAATAATTTCAAGGCTGTGGCAATGCAACACAGACCTGAC 20798
QY 903 TGAAGTAAAGTTTGAAGAAATGTTT-----GAGAACAAAGGTTTAAAGAG----- 952
DB 20799 GCTGTTCAAGTTCTGAAGAGCTCTGATCTCTCAAGGCCCCAGAGTTTCAAGAGAGAT 20858
QY 953 -----GTGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1006
DB 20859 CTGTAGGCTGATTAAG 20918
QY 1007 CTAATGAGAAATCAAAATGAGATGTTTCAATGATTTAAGAGCCCTGTTAAGGTTGA 1066
DB 20919 TGCCCCAG---GACTGCTGGGTACTTACCGAAGAGAGCCGCTCAGTTGGCAGCTGGA 20975
QY 1067 AGACATTAACATTAATCTGACCAATCGGGGAGCTTTTATATGTAATGCTGGCAATT 1126
DB 20976 TGAAGCTAGGCTGCTGACGCCAGTGAAGCTTCTGCTCTCCAGTGTGATGCTCACTG 21035
QY 1127 TAAATGAGAGAGAGAGATGATGACAGTTGATGATGATGATGATGATGATGATGATGAT 1186
DB 21036 CATGGGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 21095
QY 1187 CATGTAAG 1246
DB 21096 CCGGCTCAATCCGGGCGCTCTACCCAGCTGCTGCTCAATGATGATGATGATGATGATGATGAT 21155
QY 1247 ATGCTGTCGAGAGAGAGATTA-----GTGAGTCTAATGAGGAGAGAGAGAGAGAG 1297

QY 1297 ATGCTGTCGAGAGAGAGATTA-----GTGAGTCTAATGAGGAGAGAGAGAGAGAG 1297

Query Match 4.8%; Score 296.8; DB 4; Length 33769;
Best Local Similarity 55.3%; Pred. No. 5.6e-60;
Matches 823; Conservative 0; Mismatches 622; Indels 43; Gaps 11;

DB 21156 AGGCGAAGAGACACCTTAACACAGATCTGTTTGTGTTTGTGTCATAGTGAAC 21215
QY 1298 AAGATTAAGCAGCTTGTGTAGTTATGTAAGAGAGTGAAGGCTTCAAGACCTACAGTGTG 1357
DB 21216 ATGAAACAAACAGACAGTGAACCTTCCCTTAATCTCACCACATGGAAATTAACACAGAGT 21275
QY 1358 TTGATTAAATAGAAATGATTGGAGAAAGATTAACGTGAGAGAGTGAATTTTCAAGC 1417
DB 21276 ATGGTTTCAGATATCTGCTTACCTGGCAATCAAAATATATCTATTTCCAACTGTGTT 21335
QY 1418 TTGAGTACCTCTACATACAGACACTGTGTAAATGCTTCAAGACATGATCCCTGCC 1477
DB 21336 CTCAGTTAAACAGTAAATCTGGGACCTTCCCTTGTGATGAAGAAATCTTGTTTTC 21395
QY 1478 TCAAGGACCTTACAGCCAAACAAAGATTAAGAAATACACACCAATAC---TATTATA 1533
DB 21396 TTTTGAATGATGCTAGTGAATCTGCTGTAAATTTTAAAGAACCTTCAAGTTATTTCT 21455
QY 1534 GGACACTTGTGTGAATATCAAGAAAGAAATACGATCTAGTACTGTAGATGTCAAC--- 1590
DB 21456 GATTTTGTGCTACCAAGAAATGCTGTAAATGAACTCTAAAGGCAATTCAAACACT 21515
QY 1591 --GGCATCAAGATATCTTCTAGTTTCAAGAGTTTCAAGTCGGCGGGCGGGTGGCTC 1648
DB 21516 CAGGATGAATATTTATTTAGTATTAAGAAATGAGCTATCGGCTGGGCCCTGAGCTC 21575
QY 1649 AGCCTGTATATCCACACACTTTGGGAGGCGGAGCGGTGATCAAGGTCAAGATC 1708
DB 21576 AACCTCTAATCCACACACTTTGGGAGGCGGAGCGGTGATCAAGAGTCCGGAGATC 21635
QY 1709 AAGACCATCTGTGTTAAACAGGTGAACCCCGTCTCTCAAAATAATTAATAATTAAGCC 1768
DB 21636 AAGACCATCTGTGTTAAACAGGTGAACCCCGTCTCTCAAAATAATTAATAATTAAGCC 21695
QY 1769 AAGCGGTGTGGGGGCGGCTGTAGTCCAGCTACTCAGAGGCTGAGGAGGAATGGC 1828
DB 21696 AAGCGGTGTGGGGGCGGCTGTAGTCCAGCTACTCAGAGGCTGAGGAGGAATTCAT 21755
QY 1829 GTGAACCCCGGAGGTAGATTGGC--GTGAGCCGAGATCGCGCACTGCGCTCAGCGTGG 1887
DB 21756 TTGAACCCCGGAGGTAGATTGGC--GTGAGCCGAGATCGCGCACTGCGCTCAGCGTGG 21815
QY 1888 GCGACAGAGTGAAGTCTGCTTCAAAAAAAAAAAAAAAAAAAG 1935
DB 21816 GCGACAGAGTGAAGTCTGCTTCAAAAAAAAAAAAAAAAAAAG 21863

RESULT 14
US-09-543-771-8
; Sequence 8, Application US/09543771
; Patent No. 6780609
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/09/543,771
; EARLIER FILING DATE: 2000-04-05
; EARLIER APPLICATION NUMBER: US 09/229,319
; EARLIER FILING DATE: 1999-01-13
; EARLIER APPLICATION NUMBER: US 60/071,449
; EARLIER FILING DATE: 1998-01-13
; EARLIER APPLICATION NUMBER: US 60/105,511
; EARLIER FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 62
; SEQ ID NO 8
; LENGTH: 33769
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (33739) (33749) (33758)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unkno

US-09-543-771-8
Query Match 4.8%; Score 296.8; DB 4; Length 33769;
Best Local Similarity 55.3%; Pred. No. 5,6e-60;
Matches 823; Conservative 0; Mismatches 622; Indels 43; Gaps 11;
QY 486 GCTGGCGGGGTGGCTCAAGCCCTGTAAATCCAGCCCTTTGGAGGCGCAGAGTGGAGAT 545
DB 20381 GCCGGGTGGGGTGGCTCAAGCCCTGTAAATCCAGCACTTGGAGGCGGAGGAGATCC 20440
QY 546 CATGAGTCAAGAGTTTGAACCAAGCTTCAAGTGTGAACCCCGTCTTACTATAA 605
DB 20441 CTTGAGCTTAAAGATTGAGACCAAGCTTCAAGCAAT--GTGAGCCCGCTCTTACCAA 20499
QY 606 GATACAAAAAATTTAGCAGGTGTGGGCAACACTGTAACTCCCACTCTTGAAGAG 665
DB 20500 AATTAAC--AAAAATTAAGCAAGTGTGGGCACTCTGTAGTCCAGCTTACCTGGAGGC 20558
QY 666 TGAGGAGAGAAATGCTTGAACCCAGAGGAGAGG--TGCAGCGCAAAAGATCAAC 724
DB 20559 TGAGGAGAGAAATGCTTGAACCCAGAGGAGAGG--TGCAGCGCAAAAGATTAAGAC 20618
QY 725 AATGCACTGATGCTGATGACAGGCAAGCTCCAACTCAAAAAA---AAAAAA 782
DB 20619 AATGCACTGATGCTGATGACAGGCAAGCTCCAACTCAAAAAA---AAAAAA 20678
QY 783 GAAATATGTAGTGTGTCTATTAACAAATTAATAGCAATGAGAGCAAAAGTCTAAAGCC 842
DB 20679 AAAAAAATGTCATAGGTGTGGGCTACAGATATGCTTTTGTGCTTACTTAAGATGAAC 20738
QY 843 TATGACATGTATGATCAAGAAATCTGTGGAAACATAATTAAGGAACTTAACCCAGTCC 902
DB 20739 GTGCCAATTTCTATTAAGAAATTTCAAGGCTGTGGCAATGCCACAGACCTTGAC 20798
QY 903 TGAAGTAAAGTTTGAAGAAATGTTT---GAGCAAAAGGTTTAAAGAA--- 952
DB 20799 GCTGTTCAGATTTCTGAAGATCTCTGATCTTCTCAGAGGCCCAAGATTTCAGAGAAAGT 20858
QY 953 -----GTGAAAAAATAATTAATAATCAAGTTTGTGTGAGAAATGGATAGGAG 1006
DB 20859 CTGTAGGCTGTAGTAAAGAAAGCCCTTCAAAAGCCCTGGGAGCAAAAGGAGAAAGGG 20918
QY 1007 CTAACTAGAGAAATCAATAGAAATGTTTCAATGTATGTTAAAGACCTTGTAGAGGTGA 1066
DB 20919 TGCCTCAG---GACTGCGGTGTACTTACCGGAGGAGCGTCCAGTTTGGACGCTGGA 20975
QY 1067 AGACCATTAATATCTGACCATCGCGGACCTTTTATTTTATGTAATGCTTGGCAATT 1126
DB 20976 TGAAGTGAAGTTGGGCTGAGCCAGTGAAGCTTCTCTCCAGGTGATGTATGTC 21035
QY 1127 TAAATAGAGAGCAGAAATGTAGACAGTTGATGAGTCAAGATTGAATTTCTGCCAGA 1186
DB 21036 CATGGGCGCAGTAAATGTCCGAGTCCAAATGATCTTCGCGGTGCTGCATCATCTCTG 21095
QY 1187 CATGGAAGAGAGACAGTGAAGGAGGAGTTTCAAGAGATTTCAAGACAGAAAGTTA 1246
DB 21096 CCCGCTCAATCCGGGGCGCTCACCCCACTGTCTCAAGTACATGTAACCTGTACGGGGC 21155
QY 1247 ATGTGCTGCGCAGTGGCATCTA-----GTGAGCTATCTGAGGAGAAAGAGTG 1297
DB 21156 AAGGCAAGAGAGCAGCTTAACCAAGATCTGTTTGTGTTTGTCTGCTATGATGAGAC 21215
QY 1298 AAGATTAAGCAGCTTGTGTAGTTATGTAAGAGAGTGAAGGCTTCAAGACCTACAGTGTG 1357
DB 21216 ATGAAACAAACAGACAGTGAACCTTCCCTTAATCTCACCACATGGAAATTAACACAGGT 21275
QY 1358 TTGATTAAATAGAAATGATTGGAGAAAGATTAACGTGAGAGAGTGAATTTTCAAGC 1417
DB 21276 ATGGTTTCAGATATCTGCTTACCTGGCAATCAAAATATATCTATTTCCAACTGTGTT 21335
QY 1418 TTGAGTACCTCTACATACAGACACTGTGTAAATGCTTCAAGACATGATCCCTGCC 1477
DB 21336 CTCAGTTAAACAGTAAATCTGGGACCTTCCCTTGTGATGAAGAAATCTTGTTTTC 21395

QY 1478 TCAAGGACTTACAGCCAAAAACAAGATAGAAATACACACCAATAC-----TATTATA 1533
Db 21396 TTTTGTATGATGACCTAGTATCTCTGCTGTAAAGTTTAAAGAACTTAGGTTATTTCT 21455
QY 1534 GGACACTTGTGTGAATATCAAGAAAGAAATACATTAAGTACTGTAGTGTGACAC--- 1590
Db 21456 GATTTTGTCTACCATGAAATGCTGTAAATGAACCTTAAGGCAATTCAAACACT 21515
QY 1591 --GGCATCAAGAATATCTTCTAGTTTCAGAAAGTTTCAGATCGGCCGCGGCTGCTC 1648
Db 21516 CAGGATGAAATATATTATTAGGTATTAAGAAATGACTATTCGGCTGGCCCACTGCTC 21575
QY 1649 ACCGCTTATATCCAGACACTTTGGAGAGCCGAGCGGCTGATCAAGGTCAGAGATC 1708
Db 21576 ACACCTTAAATCCAGACACTTTGGAGAGCCGAGCGGCTGATCAAGGTCAGAGATC 21635
QY 1709 AAGACATCTCTGTTAAACAGGTGAACCCCGTCTCTACAAAATAATTAATAATTAGCC 1768
Db 21636 AAGACATCTCTGTTAAACAGGTGAACCCCGTCTCTACAAAATAATTAATAATTAGCC 21695
QY 1769 AGCGGTGTGTGCGGCGGCTGTAGTCCAGCTACTCAGAGGCTGAGGAGCAAGATGCG 1828
Db 21696 AGCGGTGTGTGAGTACGACCTGTAGTCCAGCTACTTAGAGGCTGAGGAGCAAGATCAT 21755
QY 1829 GTGAACCCCGGAGGTAGAGTTTGC-GTGAAGCCAGATCGCGCTCAAGCTTCAAGCTGG 1887
Db 21756 TTGAACCCCGGAGGTAGAGTTTGC-GTGAAGCCAGATCGCGCTCAAGCTTCAAGCTGG 21815
QY 1888 GCGACAGAGTGAAGTGTCTCAAAAAAAG 1935
Db 21816 GCGACAGAGTGAAGTGTCTCAAAAAAAG 21863

RESULT 15
US-09-791-211-3
Sequence 3, Application US/09791211
Patent No. 6448080
GENERAL INFORMATION:
APPLICANT: Donna T. Ward
APPLICANT: Andrew T. Walt
TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
FILE REFERENCE: RTS-0205
CURRENT APPLICATION NUMBER: US/09/791,211
CURRENT FILING DATE: 2001-02-23
NUMBER OF SEQ ID NOS: 90
SEQ ID NO 3
LENGTH: 87543
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: 7421
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 7427
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 11609
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 12605
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 12742
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29370
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29422
OTHER INFORMATION: unknown
NAME/KEY: unsure

LOCATION: 29979
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29980
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29981
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 30136
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 30140
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31205
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31206
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31592
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 33095
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 33160
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 34066
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 34072
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 36816
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 39020
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 42164
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 42459
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46808
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46823
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46826
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 47291
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 52786
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 52787
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 53384
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 54684
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59215

OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59235
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59242
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 63290
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 66614
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68660
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68697
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68718
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68733
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68739
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 69785
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 79134
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 79198
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 86336
OTHER INFORMATION: unknown
OTHER INFORMATION: unknown
US-09-791-211-3

Query Match 4.7%; Score 291.6; DB 4; Length 87543;
Best Local Similarity 52.0%; Pred. No. 1.5e-58;
Matches 769; Conservative 0; Mismatches 704; Indels 5; Gaps 5;

458 TTGTACATTTGGAAATATATATAGTTGTTACCTGCGCCGCTGCTCAAGCTGTAATCCCA 517
9124 TGGAAATATTTCTAAAAAATATATATGTTGGCCGGGTGCACTGCTCATGCTGTAATCCCA 9183
518 GCCCTTTGGAGAGCCAGAGTGGGACAGATCATGAGTCAAGAGTTTGAACAAGCTAGCC 577
9184 GCACTTTGGAGAGCCAGAGCGGGGTGATCCAGAGTCAAGAGTCAAGAGTCAAGAGTCC 9243
578 AACATGTTGAAACCCCGTCTCTAATAAGATCAAAAAATTAGCCAGGTGTGTCGCA 637
9244 AACATGTTGAAACCCCGTCTCTAATAAGATCAAAAAATTAGCCAGGTGTGTCGCA 9303
638 CACCTGTAAACCCCGTCTCTAATAAGATCAAAAAATTAGCCAGGTGTGTCGCA 697
9304 TCCCTGTAAACCCCGTCTCTAATAAGATCAAAAAATTAGCCAGGTGTGTCGCA 9363
698 AGAG-GTGCAGGAGCAAGATCAACCAATGCACTGTAGCTGTAGTGAAGGCAAGAC 756
9364 GAGAGTGTGAGAGGAGGAGAGTGTGCTCACTGCACTCCAGCTGCGGCAACAGAGCAAC 9423
757 TCCAACTCAAAAAAAT 816
9424 TCTGTCTCAAAAAAATCAACCAATGCACTGTAGTGAAGTCAAGATCTGTGAGCAAC 9483
817 GAGAGTGTGAGAGGAGGAGTGTGCTCACTGCACTGTAGTGAAGTCAAGATCTGTGAGCA 876

9484 AAACATATGATTTAGACATGGGTTGGCAAACTATAGCCCTGTAGGCCAATCTGATGCTG 9543
877 CATTAATAGGGAACCTAACCCAGCTCTGGAAGTATGTTGGAAAGAAATGTTGAGA 936
9544 TTTTATTTTTTTTATTTTTTTGACATAGGGTCACTACAGGCTGTGACACAGGCTGGAGAG 9603
937 CAAGGGTTAAAGAGTGAATAAAAAATTAATAATACAGTTTGTAGCTGTGAGAAAG 996
9604 CAGTGTATGATCATAGCTCAGTGTAACTCAAAATTCCTGGGCTCAAGCAATCTCTGCG 9663
997 GGATAGGAGCTAACTAGAGAAATCAAAATAGGAATGTTTCATGTATGTAAAGACCTG 1056
9664 CTCACCTAGCTTCCCAATAGCTACAGGCACTACAGCACTACCAAGCTTATTAAGA 9723
1057 GTAAGGTTGAAGACCATTAATATCTGACCAATCGGGGACTTTTTTTTATGTAATG 1116
9724 AATTTTTTTTGTGAGACAGTCTCAGATATGTGCGGAGCTGTTTTCAAACTGCTTG 9783
1117 CTTGGCAATTTAAATAGAGAGACAGAAATGTAGACAGTTGATGTAGCAAGTTGAAG 1176
9784 CCTCAATCAGCTCTCTTCTTCAAGCTCTTAAAGTGTGGATTTATAGGCTGAGGCAATC 9843
1177 TTCTGCCAGACATGTGAAGAGAGAGACAGTATGAGGCAAGATTGAAGATTTATCAAG 1236
9844 ACGCTTGACATATGTTTTTTTAAATATTAAGTTTCTCAGAACACAGCAATGCTTTGTT 9903
1237 ACAGAAATTAATGTGTGCGCAAGTGCATCTAGCTGATCTAATCTGAGGAAAGAGT 1296
9904 ATGTGTATGTAGGGCT-GCTGAGTTAAGTGTGTTGCTTAAAGGCTTATGAGCTTAT 9962
1297 GAAGATAGCACTGCTGATAGTATGAAGAGTGTGAAGGCTTCAAGAGCTTACAGGT 1356
9963 AAAGCTGAAATCTATCTATCTGCTCTTATGAAGTGTGTTCTGACCTGTATCTAG 10022
1357 GTTGAATTAATGAAGATGATTTGGAAAGAAATACGTGAGAGATGAGATTTTCAGG 1416
10023 ACTAGCTGTCTCAAAATTTCTCAATGAATTTGGAAGTTTCTCAGCAATTTTCTGACC 10082
1417 CTTGAGTGCATCTCAATACAGACATGTGTCTAAATGCTTCAAGACATGATCTCTGCC 1476
10083 AATAATGCACTTGAAGTGAAGTAAATTAAGCAGATTAACAACAAATCTCATGATTTGG 10142
1477 CTCAGAGGACTTACAGCCAAAAACAAGATTAAGAAATCAACCAATCTATTAATAGA 1536
10143 AATTTAAATTAACCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10202
1537 CACTGTGTAGAAATTAAGAAAGAAATTAAGATTAATTAATTAATTAATTAATTAATTA 1596
10203 AAAAATTTTAACTACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10262
1597 AAAGATATCTTCTAGTTTCAGAAAGTTTCAGATCGGCGGCGGCTGCTCAGGCTGT 1656
10263 AATAGTTAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10322
1657 AATCCAGCACTTTGGAGAGCGGAGGCGGAGTGTATCAAGGTCAGAGATTAAGACCT 1716
10323 AATCCAGCACTTTGGAGAGCGGAGGCGGAGTGTATCAAGGTCAGAGATTAAGACCT 10382
1717 CCTGTTTAAACAGGTTGAACCCCGTCTCTAATAAAATTAATAAAATTAAGCAAGGCTG 1776
10383 CTTGGCAACATGTGTAAATTCCTGTTCTTCTAATAAAATTAATAAAATTAAGCAAGGCTG 10441
1777 TGGCGGCGCTGTAGTCCCACTACTCAGAGGCTGAGAGCAGAGAAATGCGGTGAACC 1836
10442 TGGCAAGGCGCTGTAGTCCCACTACTCAGAGGCTGAGAGCAGAGAAATGCGGTGAACC 10501
1837 GGAAGTATAGTTTG-CGTGAGCGAGATCGGCGCACTGCGCTTCAAGCTGCGGCGCA 1895
10502 GGAAGCGGAGGTTGCGGTGAGCGAGATTTGGCCCATGCACTTCCAGCT-GGCGACAG 10560
1896 GTGAGACTGCTCTCAAAAAAATTAATAAAATTAATAAAATTAATAAAATTAATAAAAT 1933
10561 GCGACACTTGTCTCAAAAAAATTAATAAAATTAATAAAATTAATAAAATTAATAAAAT 10598

Tue Dec 28 15:07:57 2004

us-10-786-065-3_copy_20000_26200.rn1

Page 20

Search completed: December 27, 2004, 11:08:24
Job time : 361.254 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2004, 17:48:22 ; Search time 1996.28 Seconds
(without alignments)
17296.376 Million cell updates/sec

Title: US-10-786-065-3_COPY_20000_26200

Perfect score: 6201
Sequence: 1 caccatgggctcaggcaatt.....aagagtaaggaagctgtg 6201

Scoring table:
IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4105333 seqs, 2784095677 residues

Total number of hits satisfying chosen parameters: 8210666

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6201	100.0	53332	9	US-09-801-861-3
2	6201	100.0	53332	14	US-10-224-562-3
3	452	7.3	454	13	US-10-027-632-4668
4	452	7.3	454	15	US-10-027-632-4668
5	402.6	6.5	2403	17	US-10-323-281-751
6	402.6	6.5	3802	10	US-09-814-353-20512
7	375	6.0	75252	13	US-10-087-192-904
8	373.6	6.0	73145	13	US-10-087-192-274
9	365.2	5.9	268685	15	US-10-265-071-22
10	365.2	5.9	268685	15	US-10-025-966A-72
11	338	5.5	2824	10	US-09-814-353-20332
12	338	5.5	91000	15	US-10-002-491-10

13	336	5.4	52242	16	US-10-052-482-172	Sequence 172, App
14	334.4	5.4	31474	10	US-09-764-891-8149	Sequence 8149, Ap
15	334	5.4	111084	18	US-10-723-860-1627	Sequence 1627, Ap
16	333.4	5.4	104245	15	US-10-160-807-4	Sequence 4, Appl1
17	333.4	5.4	104245	16	US-10-655-847-4	Sequence 4, Appl1
18	333.4	5.4	170245	17	US-10-717-597-322	Sequence 322, App
19	331	5.3	506	14	US-10-066-543-1709	Sequence 1709, Ap
20	331	5.3	549	14	US-10-066-543-194	Sequence 194, App
21	330.6	5.3	164875	15	US-10-085-117-322	Sequence 322, App
22	327.6	5.3	107745	17	US-10-322-281-368	Sequence 268, App
23	327.6	5.3	128779	15	US-10-081-337-358	Sequence 38, Appl
24	323.6	5.2	561515	17	US-10-741-601-5682	Sequence 5682, Ap
25	323.6	5.2	21371	10	US-09-764-891-1018	Sequence 7018, Ap
26	323.4	5.2	350570	18	US-10-417-375-146	Sequence 146, App
27	323.2	5.2	136726	15	US-10-085-117-244	Sequence 244, App
28	321.6	5.2	1503841	9	US-09-795-668-1	Sequence 1, Appl1
29	321.6	5.2	1503841	9	US-09-795-668-1	Sequence 1, Appl1
30	321.6	5.2	1503841	9	US-09-946-807-1	Sequence 1, Appl1
31	320.8	5.2	32248	9	US-09-764-864-1769	Sequence 1769, Ap
32	320.8	5.2	32248	9	US-09-764-877-3487	Sequence 3487, Ap
33	320.8	5.2	32248	16	US-10-242-515-3487	Sequence 3487, Ap
34	319.8	5.2	143239	18	US-10-723-860-546	Sequence 546, App
35	319	5.1	16554	15	US-10-017-161-1961	Sequence 1961, App
36	319	5.1	16554	15	US-10-292-798-1609	Sequence 1609, Ap
37	319	5.1	37571	13	US-10-087-192-1516	Sequence 1516, Ap
38	317.6	5.1	304505	16	US-10-271-416-1	Sequence 1, Appl1
39	317.2	5.1	94720	17	US-10-052-482-160	Sequence 160, App
40	316.8	5.1	41061	17	US-10-322-281-302	Sequence 302, App
41	315	5.1	96592	11	US-09-997-722-88	Sequence 88, Appl
42	314.6	5.1	209484	13	US-10-087-192-418	Sequence 418, App
43	314.6	5.1	209484	15	US-10-331-053-4	Sequence 4, Appl1
44	313.8	5.1	73857	18	US-10-085-117-310	Sequence 310, App
45	313.4	5.1	32191	15	US-10-074-024-446	Sequence 446, App

ALIGNMENTS

US-09-801-861-3	US-10-052-482-172
Sequence 3, Application US/09801861	Sequence 172, App
Patent No. US20020119544A1	Sequence 8149, Ap
GENERAL INFORMATION:	Sequence 1627, Ap
APPLICANT: YAN, Chunhua et al.	Sequence 4, Appl1
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC	Sequence 322, App
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES	Sequence 1709, Ap
TITLE OF INVENTION: THERMOF	Sequence 194, App
FILE REFERENCE: CLO01098	Sequence 322, App
CURRENT APPLICATION NUMBER: US/09/801,861	Sequence 268, App
CURRENT FILING DATE: 2001-03-09	Sequence 5682, Ap
NUMBER OF SEQ ID NOS: 10	Sequence 7018, Ap
SOFTWARE: FASTSEQ for Windows Version 4.0	Sequence 146, App
SEQ ID NO 3	Sequence 244, App
LENGTH: 53332	Sequence 1, Appl1
TYPE: DNA	Sequence 1, Appl1
ORGANISM: Human	Sequence 1769, Ap
US-09-801-861-3	Sequence 3487, Ap
Query Match	Sequence 546, App
Best Local Similarity 100.0%; Score 6201; DB 9; Length 53332;	Sequence 302, App
Mismatches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Sequence 88, Appl
	Sequence 418, App
	Sequence 4, Appl1
	Sequence 310, App
	Sequence 446, App

QY 181 ATTCAATATATCATCAAAAGTTATCTTCAAGAGCTTCAGCGCCCTAATGATGTCTAAAGAAA 240
DB 20180 ATTCAATATATCATCAAAAGTTATCTTCAAGAGCTTCAGCGCCCTAATGATGTCTAAAGAAA 20239
QY 241 ATGTGAAGAGCCCTCAGCATCTGAAGGACAGTGTTTACAGCAATTGATCAAAAAGAAAA 300
DB 20240 ATGTGAAGAGCCCTCAGCATCTGAAGGACAGTGTTTACAGCAATTGATCAAAAAGAAAA 20299
QY 301 CCACAGGCCCTTCCCTTCCCATCTGATGTAAAGAGCTTCAATTTCCATAGTAGT 360
DB 20300 CCACAGGCCCTTCCCTTCCCATCTGATGTAAAGAGCTTCAATTTCCATAGTAGT 20359
QY 361 AAATTTCTAGATACAGCTTTGAGCTGCAAGTACTGAAAGAAAGCTCCATTCAAAG 420
DB 20360 AAATTTCTAGATACAGCTTTGAGCTGCAAGTACTGAAAGAAAGCTCCATTCAAAG 20419
QY 421 GAAATTTATCTTAAATGATCTGTAATGATCTAAATTTTGTACATTTGGAATATATAGT 480
DB 20420 GAAATTTATCTTAAATGATCTGTAATGATCTAAATTTTGTACATTTGGAATATATAGT 20479
QY 481 TGTTAAGCTGCGCGGTGCTCAGCGCTGTAAATCCAGCCCTTTGGAGGCCAGATGGG 540
DB 20480 TGTTAAGCTGCGCGGTGCTCAGCGCTGTAAATCCAGCCCTTTGGAGGCCAGATGGG 20539
QY 541 CAGATCATAGAGTCAGAGATTTGAGACCAAGCTTAAAGCAATGTGTAAACCCGCTCTTA 600
DB 20540 CAGATCATAGAGTCAGAGATTTGAGACCAAGCTTAAAGCAATGTGTAAACCCGCTCTTA 20599
QY 601 CTAAAGATCAAAAATTTAGCCAGGTGTGTGGCGACACTGTAAACCCAGCTGTCTGA 660
DB 20600 CTAAAGATCAAAAATTTAGCCAGGTGTGTGGCGACAACCTGTAAACCCAGCTGTCTGA 20659
QY 661 GAGAGTGAGCAGAGAAATTCCTTGAACCCAGAGGCAAGGTGCAAGCAAGCAAGATCA 720
DB 20660 GAGAGTGAGCAGAGAAATTCCTTGAACCCAGAGGCAAGGTGCAAGCAAGCAAGATCA 20719
QY 721 CACCAATGCACTGTAGCTGTGATGACAGGCAAGATCTCCAACTCAAAAAAAAAAAAAA 780
DB 20720 CACCAATGCACTGTAGCTGTGATGACAGGCAAGATCTCCAACTCAAAAAAAAAAAAAA 20779
QY 781 AAGAAATATGTAAAGTTGTCTATAACAATTAATGAGCAGTGAAGCAAAAGCTTAAG 840
DB 20780 AAGAAATATGTAAAGTTGTCTATAACAATTAATGAGCAGTGAAGCAAAAGCTTAAG 20839
QY 841 CCTATGACCATGTGTAACTAGAAATCTGTGGGAAACATPATTAAGGAACTTAACCCAGT 900
DB 20840 CCTATGACCATGTGTAACTAGAAATCTGTGGGAAACATPATTAAGGAACTTAACCCAGT 20899
QY 901 CCTGGAAGTAAGTTTGAAGAAATGTTTGAAGCAAAAGGTTTAAAGAGTGAAGAAA 960
DB 20900 CCTGGAAGTAAGTTTGAAGAAATGTTTGAAGCAAAAGGTTTAAAGAGTGAAGAAA 20959
QY 961 AAAAAATTAATAACAGTTAGCTGTGTGAGAAATGGATAGGAGCTTAATAGAGAAAT 1020
DB 20960 AAAAAATTAATAACAGTTAGCTGTGTGAGAAATGGATAGGAGCTTAATAGAGAAAT 21019
QY 1021 CAAATAGGAATGTTTATGTATGTTAAGAACCTGTGAAGGCTGAAGCAATTAATTA 1080
DB 21020 CAAATAGGAATGTTTATGTATGTTAAGAACCTGTGAAGGCTGAAGCAATTAATTA 21079
QY 1081 TCTGCAACCATGCGGGGACTTTTATGTATGTTAAGCTTGAAGCTTAAATAGAGAGCA 1140
DB 21080 TCTGCAACCATGCGGGGACTTTTATGTATGTTAAGCTTGAAGCTTAAATAGAGAGCA 21139
QY 1141 GAGAAATGTAGACAGTTGATGATGACAGATGTAAGTTCTGCCAGCATGTGAAGAGAG 1200
DB 21140 GAGAAATGTAGACAGTTGATGATGACAGATGTAAGTTCTGCCAGCATGTGAAGAGAG 21199
QY 1201 AGACAGTGAAGCAAGAGATTTGAAGATTAATCAAGACAGAAAGTTAATGTGCTGGCAGT 1260
DB 21200 AGACAGTGAAGCAAGAGATTTGAAGATTAATCAAGACAGAAAGTTAATGTGCTGGCAGT 21259

QY 1261 GGCATCTAGTCTGAGTCTAATCTGAGGGAAGAGTAAGTAATAGCAGCTTGTGATAGT 1320
DB 21260 GGCATCTAGTCTGAGTCTAATCTGAGGGAAGAGTAAGTAATAGCAGCTTGTGATAGT 21319
QY 1321 TATGAAGAGAGTGAAGGCTTCAAGGACCTTACAGGCTTGATTAATAGAAAGATTTG 1380
DB 21320 TATGAAGAGAGTGAAGGCTTCAAGGACCTTACAGGCTTGATTAATAGAAAGATTTG 21379
QY 1381 GAGAAAGATTAACCTGTAGAGAGTGAAGATTTTCAAGCTTGAAGCTCTCAATACAGA 1440
DB 21380 GAGAAAGATTAACCTGTAGAGAGTGAAGATTTTCAAGCTTGAAGCTCTCAATACAGA 21439
QY 1441 CACTGTGCTAATGCTTCAAAAGACATGATCCCTGCCCTCAAGGACTTACACCCAAAAAC 1500
DB 21440 CACTGTGCTAATGCTTCAAAAGACATGATCCCTGCCCTCAAGGACTTACACCCAAAAAC 21499
QY 1501 AAGAGATTAAGAAATCACACCAATCTATTATAGGACACTTGTAGAAATATCAAGAAAG 1560
DB 21500 AAGAGATTAAGAAATCACACCAATCTATTATAGGACACTTGTAGAAATATCAAGAAAG 21559
QY 1561 AAATACGATCTAGTACTGTAGATGTGCAACGGCATCAAAAGATATCTTCTAGTTTCAAGA 1620
DB 21560 AAATACGATCTAGTACTGTAGATGTGCAACGGCATCAAAAGATATCTTCTAGTTTCAAGA 21619
QY 1621 GTTTCAATCGGCGGGCGGGTGGCTCAGCGCTGTAAATCCAGACCTTTGGAGGCCGA 1680
DB 21620 GTTTCAATCGGCGGGCGGGTGGCTCAGCGCTGTAAATCCAGACCTTTGGAGGCCGA 21679
QY 1681 GCGGGGTGATCACAAAGTCAGAGATCAAGACCAATCTGTGTAAACACGGTGAACCCCG 1740
DB 21680 GCGGGGTGATCACAAAGTCAGAGATCAAGACCAATCTGTGTAAACACGGTGAACCCCG 21739
QY 1741 TCTCTACAAAAATATTAATAATTAACCAAGCGTGTGTGGCGGCGCTGTAGTCCAGCT 1800
DB 21740 TCTCTACAAAAATATTAATAATTAACCAAGCGTGTGTGGCGGCGCTGTAGTCCAGCT 21799
QY 1801 ACTCAGAGGCTGAGGCGAGAAATGGCGTGAACCCGGAGGTTAGATTTGGTGAGCCG 1860
DB 21800 ACTCAGAGGCTGAGGCGAGAAATGGCGTGAACCCGGAGGTTAGATTTGGTGAGCCG 21859
QY 1861 AGATCGCGCACTGCGCTCAGACCTGCGGCAAGAGTGAAGCTGCGTCAAAAAAAAAA 1920
DB 21860 AGATCGCGCACTGCGCTCAGACCTGCGGCAAGAGTGAAGCTGCGTCAAAAAAAAAA 21919
QY 1921 AAAAAAAAAAAAAAGTTTCAATCTTAAACACATCTGATTTCAACGTTTGAATAGGAG 1980
DB 21920 AAAAAAAAAAAAAAGTTTCAATCTTAAACACATCTGATTTCAACGTTTGAATAGGAG 21979
QY 1981 AGCATGTTACAGGAGAGAAATGTTTCAAGCAAGGTTACAGATAGGGAATTAAGGAT 2040
DB 21980 AGCATGTTACAGGAGAGAAATGTTTCAAGCAAGGTTACAGATAGGGAATTAAGGAT 22039
QY 2041 ATGTTCAAGGAAGAGACCCCAAGATCAATGTTTGTAGGTTTGAAGGAAACACAGTGT 2100
DB 22040 ATGTTCAAGGAAGAGACCCCAAGATCAATGTTTGTAGGTTTGAAGGAAACACAGTGT 22099
QY 2101 TTGCAATCTCCAGGTTCCATTAATGCTGTATGAATCAATATGTGTGTTAGCAACTGCA 2160
DB 22100 TTGCAATCTCCAGGTTCCATTAATGCTGTATGAATCAATATGTGTGTTAGCAACTGCA 22159
QY 2161 TTTTAAAAAATGAATTAATGATGAGAAGAGAAATGAATTAATAGCATGATTAACAT 2220
DB 22160 TTTTAAAAAATGAATTAATGATGAGAAGAGAAATGAATTAATAGCATGATTAACAT 22219
QY 2221 TTGAAGAGCAAGATTAATTTTCTGCAACTTTTGTCTCAATTTGAATCTTATATTT 2280
DB 22220 TTGAAGAGCAAGATTAATTTTCTGCAACTTTTGTCTCAATTTGAATCTTATATTT 22279
QY 2281 TTATGTATGATGTGAATTAACAGATCAATATATTTCTTACTGTAGACTGTGACAAA 2340
DB 22280 TTATGTATGATGTGAATTAACAGATCAATATATTTCTTACTGTAGACTGTGACAAA 22339
QY 2341 AATCTTTAAGACATGCGCTGTGTAACTCTTATTTTGAAGAGAGAAATCCAAAGATC 2400

22340 AATCTTAAAGCATGCGCTGGCTTAATCTCTATTGTCAGAGGAAATCCAAAGATC 22399
2401 TGAAGAGCAAAATTTTGCCTGAGGTATAGAAACAGCTTATGCCATTGCTAAAAGTA 2460
22400 TGAAGAGCAAAATTTTGCCTGAGGTATAGAAACAGCTTATGCCATTGCTAAAAGTA 22459
2461 TTCTTAGTAAATTTCTTCCCACTAGTGCATATGCACTTCTAGTTCTGTGGCTGA 2520
22460 TTCTTAGTAAATTTCTTCCCACTAGTGCATATGCACTTCTAGTTCTGTGGCTGA 22519
2521 AATPACGAATATATTAAGTAAACAGATATCACAAGCTGGGGAATATATTTGGGTAGTG 2580
22520 AATPACGAATATATTAAGTAAACAGATATCACAAGCTGGGGAATATATTTGGGTAGTG 22579
2581 GCTGAGAGCTCATTTTCTAAGAAATGTGGAACCTTAGCGAGGTATGTGTGCTCACACT 2640
22580 GCTGAGAGCTCATTTTCTAAGAAATGTGGAACCTTAGCGAGGTATGTGTGCTCACACT 22639
2641 AATPATTCCAGCACTTTGGAGGCTCAAGATCAAGAAATCGCTTGAACCCAAAGATTCAAG 2700
22640 AATPATTCCAGCACTTTGGAGGCTCAAGATCAAGAAATCGCTTGAACCCAAAGATTCAAG 22699
2701 CTAGCACTGGGCAACATAGGAAGCTCATCTCTACAAAATTTTAAATTCAGCTAGCA 2760
22700 CTAGCACTGGGCAACATAGGAAGCTCATCTCTACAAAATTTTAAATTCAGCTAGCA 22759
2761 TGGTGGCATACGCTGTAGTCCCACTACCTGGGAAGTAGTGGGTGATCGCTTGACA 2820
22760 TGGTGGCATACGCTGTAGTCCCACTACCTGGGAAGTAGTGGGTGATCGCTTGACA 22819
2821 CAGGAGTTTGAAGCTTAAGGTGAGCCATGATCACACAACTGCACTTCAAGCTTGAGTACAG 2880
22820 CAGGAGTTTGAAGGTGAGCCATGATCACACAACTGCACTTCAAGCTTGAGTACAG 22879
2881 AAGAAGACCTGCTGCTTAAATAAAGAAATGTGAAATTTTATCTTGAACAGTACAGT 2940
22880 AAGAAGACCTGCTGCTTAAATAAAGAAATGTGAAATTTTATCTTGAACAGTACAGT 22939
2941 CATTAAGCTAATTAAGTTGAGTTGAGAAATATATATGATCAGAAATTTATATCACT 3000
22940 CATTAAGCTAATTAAGTTGAGTTGAGAAATATATATGATCAGAAATTTATATCACT 22999
3001 GTGGTCTGATGATATATTAAGAAATTAAGAGACTAGAGTCAAGGATTCACCTTAAGTG 3060
23000 GTGGTCTGATGATATATTAAGAAATTAAGAGACTAGAGTCAAGGATTCACCTTAAGTG 23059
3061 TTTGTTTGTGTTTGTGAGACAGAGTCTCTTTTGTATCCAGGCTTAGATGGAATGTG 3120
23060 TTTGTTTGTGTTTGTGAGACAGAGTCTCTTTTGTATCCAGGCTTAGATGGAATGTG 23119
3121 CAGTCATGAGCTCAACGACGCTCAAACTCCAGGCTCAAAATTTCTTCCAGCTCGGCT 3180
23120 CAGTCATGAGCTCAACGACGCTCAAACTCCAGGCTCAAAATTTCTTCCAGCTCGGCT 23179
3181 CCCAAAGTCTGGAATTAAGGTGAGCCCAAGGCTTATTTGATGTGCTGGCTAGT 3240
23180 CCCAAAGTCTGGAATTAAGGTGAGCCCAAGGCTTATTTGATGTGCTGGCTAGT 23239
3241 GCTCTCAAACTTCAAGTGAAGCAAGAGTCCGGAACCTGACTCAACAGCTGTGGTT 3300
23240 GCTCTCAAACTTCAAGTGAAGCAAGAGTCCGGAACCTGACTCAACAGCTGTGGTT 23299
3301 TAAGCTGAGGCTCTGATCTTCAACATGAGTCAAGCTGAGTGTGTGCTGACAAG 3360
23300 TAAGCTGAGGCTCTGATCTTCAACATGAGTCAAGCTGAGTGTGTGCTGACAAG 23359
3361 GCCAAGCACTGAGCAAGAGATCTAAGTTAGCAATTTAGTAATCAAGGTTGATATGTGT 3420
23360 GCCAAGCACTGAGCAAGAGATCTAAGTTAGCAATTTAGTAATCAAGGTTGATATGTGT 23419
3421 AATGACAAATPAGAAATGGAAGAAATGTGAATACTAATCAAAAGAGATTCACCTTGG 3480

23420 AATGACAAATPAGAAATGGAAGAAATGTGAATACTAATCAAAAGAGATTCACCTTGG 23479
3481 GTAATGTGACATGAGAGGGAAGAGATGGGCCAAACATTAATCTGTTTGTGTTGACT 3540
23480 GTAATGTGACATGAGAGGGAAGAGATGGGCCAAACATTAATCTGTTTGTGTTGACT 23539
3541 GAGAGGAAATTTGATGCTCTATTAAACAGAAATPAGAGAAAGAGTTGTTGAGAGAA 3600
23540 GAGAGGAAATTTGATGCTCTATTAAACAGAAATPAGAGAAAGAGTTGTTGAGAGAA 23599
3601 GAGAGGCTCTGTTTCAAGCTGTGAGTCCAGGAGTACAGATCTTCAAGGGAAT 3660
23600 GAGAGGCTCTGTTTCAAGCTGTGAGTCCAGGAGTACAGATCTTCAAGGGAAT 23659
3661 GAGAGTGAAGCACTTAAAGGAAATCTGTGCTCAGAAAGGAGCTGTGAGCTGTGT 3720
23660 GAGAGTGAAGCACTTAAAGGAAATCTGTGCTCAGAAAGGAGCTGTGAGCTGTGT 23719
3721 AGATCTGAGGCTCATCAGACATAGATTTAGAAACAGAGATAGGCAACCAAAAGAGC 3780
23720 AGATCTGAGGCTCATCAGACATAGATTTAGAAACAGAGATAGGCAACCAAAAGAGC 23779
3781 AATATACAAAGAGAGAGAGATGATGAGACTTTGCTTTTATGATGAGAAAGAG 3840
23780 AATATACAAAGAGAGAGAGATGATGAGACTTTGCTTTTATGATGAGAAAGAG 23839
3841 ACAGGAAATGAAGAAATGAAGGGAAGCAGCTTGTAGAAATGTAGAGCATCTGAAAAA 3900
23840 ACAGGAAATGAAGAAATGAAGGGAAGCAGCTTGTAGAAATGTAGAGCATCTGAAAAA 23899
3901 ATACACACTGTATGGAATCAAGGGAAGAAATTTCAAGAAAGGAGGTATGTGTGACA 3960
23900 ATACACACTGTATGGAATCAAGGGAAGAAATTTCAAGAAAGGAGGTATGTGTGACA 23959
3961 GTATTACAGCATCAGGAATACAGCTTAAAGTCAATCTCTTGAATGACATTTGACTGTGG 4020
23960 GTATTACAGCATCAGGAATACAGCTTAAAGTCAATCTCTTGAATGACATTTGACTGTGG 24019
4021 ATTTGTGAGGGAACAATAATTAAGAAATTTATTTGTGGTATATGAGGCAAAAG 4080
24020 ATTTGTGAGGGAACAATAATTAAGAAATTTATTTGTGGTATATGAGGCAAAAG 24079
4081 GAAAGGTTATCCAAATCAAAACAGGTGGGAATGAGGATGCTCCCAAGGTGAGGCA 4140
24080 GAAAGGTTATCCAAATCAAAACAGGTGGGAATGAGGATGCTCCCAAGGTGAGGCA 24139
4141 TCAGTGAATGTGGGAAGGGGCAAGAGCATCCATGCGCATCCAGGCAAGCCACTCA 4200
24140 TCAGTGAATGTGGGAAGGGGCAAGAGCATCCATGCGCATCCAGGCAAGCCACTCA 24199
4201 GAAAGCTCATGATGAGTTCAAGTATCCAGAAAGTCTGTACCTTAATCTTTCTGGGTT 4260
24200 GAAAGCTCATGATGAGTTCAAGTATCCAGAAAGTCTGTACCTTAATCTTTCTGGGTT 24259
4261 TGCATAGGCTTCAATGTGTAGGCAATGATTTATTAACATATGTGGCACTGTGATCAACT 4320
24260 TGCATAGGCTTCAATGTGTAGGCAATGATTTATTAACATATGTGGCACTGTGATCAACT 24319
4321 AACCTTCAACCCCTCTCCCTCCCTTAATCATGAGCTTGTGCTTCCAGTGAACGAGTCCCTA 4380
24320 AACCTTCAACCCCTCTCCCTCCCTTAATCATGAGCTTGTGCTTCCAGTGAACGAGTCCCTA 24379
4381 TCTTAAGCTAACCAATGTGTCTGCAAGCTATCAAGTCAACTCTACAAAAGACATCACTTTG 4440
24380 TCTTAAGCTAACCAATGTGTCTGCAAGCTATCAAGTCAACTCTACAAAAGACATCACTTTG 24439
4441 GAGATTTCAAGATTTTGAAGATTTGGCTGTCAAGAAATTTAGTTGAAGATCAATATATAT 4500
24440 GAGATTTCAAGATTTTGAAGATTTGGCTGTCAAGAAATTTAGTTGAAGATCAATATATAT 24499
4501 TTCACAAATATCAAGCTGTGCTATTTTATATCAGGCGCATTAATATGTTTAAACAAAG 4560
24500 TTCACAAATATCAAGCTGTGCTATTTTATATCAGGCGCATTAATATGTTTAAACAAAG 24559

QY 4561 AGGTGATTAATTCAGATTTCTTTTAAAGCTTACAGTAGACAGTGTGTAATAG 4620
DB 24560 AGGTGATTAATTCAGATTTCTTTTAAAGCTTACAGTAGACAGTGTGTAATAG 24619
QY 4621 ATTGGGATGAGGGGCACTTTTTTTTGAATGTATATTCCTCGAACCTCTCTCTC 4680
DB 24620 ATTGGGATGAGGGGCACTTTTTTTTGAATGTATATTCCTCGAACCTCTCTCTC 24679
QY 4681 CTGTGTTCTCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4740
DB 24680 CTGTGTTCTCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 24739
QY 4741 TCCCTGAATGCTGTCTCTGTAAAGTTCAAGCTTGAAGTGAAGTAACTCAAGAC 4800
DB 24740 TCCCTGAATGCTGTCTGTAAAGTTCAAGCTTGAAGTGAAGTAACTCAAGAC 24799
QY 4801 AGTGTACAGATGTAGATGATGATGAGGAAAGTGGACAGATATGACCAATTA 4860
DB 24800 AGTGTACAGATGTAGATGATGATGAGGAAAGTGGACAGATATGACCAATTA 24859
QY 4861 GCCGGAACTGACGGGAAAAACAAGAGTCAGCGATATTTTTCTGTGATCTGAT 4920
DB 24860 GCCGGAACTGACGGGAAAAACAAGAGTCAGCGATATTTTTCTGTGATCTGAT 24919
QY 4921 ATGGAGATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4980
DB 24920 ATGGAGATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 24979
QY 4981 ATACCAATGATGTTGTTTGAAGATGATGATGATGATGATGATGATGATGAT 5040
DB 24980 ATACCAATGATGTTGTTTGAAGATGATGATGATGATGATGATGATGATGAT 25039
QY 5041 GCTCCAGACTGTGTGGGCTTAAAGTGAAGAAAGCAATCTGAGTTGAAGATTTGA 5100
DB 25040 GCTCCAGACTGTGTGGGCTTAAAGTGAAGAAAGCAATCTGAGTTGAAGATTTGA 25099
QY 5101 AATCATCAGAAATGCGGTGTTCAATTAAGACCTGTCAAGTGGGTAATGCTAAGGAG 5160
DB 25100 AATCATCAGAAATGCGGTGTTCAATTAAGACCTGTCAAGTGGGTAATGCTAAGGAG 25159
QY 5161 ATGTGTAGATGATTAACAGAAAGATCAAGACCGAACCTTAAGATTAATGTTATTA 5220
DB 25160 ATGTGTAGATGATTAACAGAAAGATCAAGACCGAACCTTAAGATTAATGTTATTA 25219
QY 5221 TTTATTAATTTATTAATGTTTATTTTAAATTTTATTTTATTTATTTATTTTAA 5280
DB 25220 TTTATTAATTTATTAATGTTTATTTTAAATTTTATTTTATTTTATTTTATTTTAA 25279
QY 5281 ACGGAGATCTGCTGTGCTGCCAGGCTGAGTGAAGTGGCGCAACTGACCTCACCTGA 5340
DB 25280 ACGGAGATCTGCTGTGCTGCCAGGCTGAGTGAAGTGGCGCAACTGACCTCACCTGA 25339
QY 5341 ACCTCGCTTCTGTGGTTCAAGGAGCTTCTGCTCAAGCTTCAAGTGTGAGACTA 5400
DB 25340 ACCTCGCTTCTGTGGTTCAAGGAGCTTCTGCTCAAGCTTCAAGTGTGAGACTA 25399
QY 5401 CAGGCAACCAACCTCACCTGATTAATTTTATTTTATTTTATTTTATTTTATTTAG 5460
DB 25400 CAGGCAACCAACCTCACCTGATTAATTTTATTTTATTTTATTTTATTTTATTTAG 25459
QY 5461 ATGTTGGCAGAGCTGTGCTTGAACCTTGAACCTTGAAGTGAACCTTGAAGCTTGC 5520
DB 25460 ATGTTGGCAGAGCTGTGCTTGAACCTTGAACCTTGAAGTGAACCTTGAAGCTTGC 25519
QY 5521 AAAAGTCTGGGATTAAGATGAGCACTGTGCTGAGCTATTTTGTGTTTATTAAGAG 5580
DB 25520 AAAAGTCTGGGATTAAGATGAGCACTGTGCTGAGCTATTTTGTGTTTATTAAGAG 25579
QY 5581 ATGGGCTTGTGATATTTGCTGAGGCTGTGCTGAGACTCTGAGACTCAAGCAATCTCT 5640
DB 25580 ATGGGCTTGTGATATTTGCTGAGGCTGTGCTGAGACTCTGAGACTCAAGCAATCTCT 25639

QY 5641 GCCTTGAGCTCTCAAGTCTGTGATTAACATGTGAGTCCCTGCGCTGGCCAGAAAT 5700
DB 25640 GCCTTGAGCTCTCAAGTCTGTGATTAACATGTGAGTCCCTGCGCTGGCCAGAAAT 25699
QY 5701 CAATATTTAGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5760
DB 25700 CAATATTTAGATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 25759
QY 5761 AAAAATTTAGAGATTTGTGAAACCAAAAGAGAGTCCCTCAGAGAGAGGATGCT 5820
DB 25760 AAAAATTTAGAGATTTGTGAAACCAAAAGAGAGTCCCTCAGAGAGAGGATGCT 25819
QY 5821 ATGATGCCAAATGCTGCAAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATA 5880
DB 25820 ATGATGCCAAATGCTGCAAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATA 25879
QY 5881 GTCATGGAACCAATGCTGCAAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATA 5940
DB 25880 GTCATGGAACCAATGCTGCAAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATA 25939
QY 5941 TATAGGGGATGATATTTGTATTTCAAGAGACAAACATTCATACATGCAAGATCTA 6000
DB 25940 TATAGGGGATGATATTTGTATTTCAAGAGACAAACATTCATACATGCAAGATCTA 25999
QY 6001 GTGTGTGACCAAGGAGATTAAGTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6060
DB 26000 GTGTGTGACCAAGGAGATTAAGTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 26059
QY 6061 GGTGTGCTGAGCTGTCTCATGACACTAATGTCAATGATGATGATGATGATGATGATGAT 6120
DB 26060 GGTGTGCTGAGCTGTCTCATGACACTAATGTCAATGATGATGATGATGATGATGATGAT 26119
QY 6121 TGCCCACTCTTCAAGTGAATGACAGAGAGGATTAAGATTAAGATTAAGATTAAGATA 6180
DB 26120 TGCCCACTCTTCAAGTGAATGACAGAGAGGATTAAGATTAAGATTAAGATTAAGATA 26179
QY 6181 GAAAGTATGAGATGTGG 6201
DB 26180 GAAAGTATGAGATGTGG 26200

RESULT 2
US-10-224-562-3
; Sequence 3, Application US/10224562
; Publication No. US20030022229A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001098DIV
; CURRENT APPLICATION NUMBER: US/10/224,562
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 53332
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-224-562-3

Query Match 100.0%; Score 6201; DB 14; Length 53332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACCATGGGCTCAGGCAATTTATGAAAGCAAAATATACAGCTTAAATAGAAATGTGAC 60
DB 20000 CACCATGGGCTCAGGCAATTTATGAAAGCAAAATATACAGCTTAAATAGAAATGTGAC 20059
QY 61 CTAATATCCAGAGAAATCTCCCTTTGTAAGATTTGTAACAAATTAATATGATGAG 120
DB 20060 CTAATATACCAGAGAAATCTCCCTTTGTAAGATTTGTAACAAATTAATATGATGAG 20119

Db 22280 TTATGTATGATGTGATACCAATACATATATATTTCTTACTGTAGACTGCAGTCAAAA 22339
QY 2341 AATCTTTAAAGCATGCGCTGTGTCTAACTTCTTATTTTGCAGAGGAAATCCAAAGATC 2400
Db 22340 AATCTTTAAAGCATGCGCTGTGTCTAACTTCTTATTTTGCAGAGGAAATCCAAAGATC 22339
QY 2401 TGAAGAGCAAACTTTGCTGAGGTTATAGAACCGAGCTTATGCCCTTATGCTTAAAGTGA 2460
Db 22400 TGAAGAGCAAACTTTGCTGAGGTTATAGAACCGAGCTTATGCCCTTATGCTTAAAGTGA 22459
QY 2461 TTTCTTAAATTTAAATTTCTTCCACTAGTGCATATGCACTTCTAGTTCTGTGGCCTGA 2520
Db 22460 TTTCTTAAATTTAAATTTCTTCCACTAGTGCATATGCACTTCTAGTTCTGTGGCCTGA 22519
QY 2521 AATACGAATATATTTATGTGAACAGCTACACAGTGTGGGAAATATATTTGGGTAGGTG 2580
Db 22520 AATACGAATATATTTATGTGAACAGCTACACAGTGTGGGAAATATATTTGGGTAGGTG 22579
QY 2581 GCTGAGAGCTCATTTTCTTAAAGAAATGTGACCTTAGGCGAGGATGTGGCTCACACCT 2640
Db 22580 GCTGAGAGCTCATTTTCTTAAAGAAATGTGACCTTAGGCGAGGATGTGGCTCACACCT 22639
QY 2641 ATAATTCAGACATTTTGGAGGCCAAGTCAAGAAATCGCTTGAACCCAAAGATTCAAGA 2700
Db 22640 ATAATTCAGACATTTTGGAGGCCAAGTCAAGAAATCGCTTGAACCCAAAGATTCAAGA 22699
QY 2701 CTATGAGTGGGCAACATAGCAAGACTCATCTCTACAAAAATTTAAATTTCAAGTGAACA 2760
Db 22700 CTATGAGTGGGCAACATAGCAAGACTCATCTCTACAAAAATTTAAATTTCAAGTGAACA 22759
QY 2761 TGTGTGCATATGCGCTGTAGTCCCACTACCTGGAAGAGTGTGGGTGATCGCTTGACA 2820
Db 22760 TGTGTGCATATGCGCTGTAGTCCCACTACCTGGAAGAGTGTGGGTGATCGCTTGACA 22819
QY 2821 CAGAGTTTGAAGGCTTAAGTGAAGCATGATCAACAATCTGACTTCAAGTTGAGTACAG 2880
Db 22820 CAGAGTTTGAAGGCTTAAGTGAAGCATGATCAACAATCTGACTTCAAGTTGAGTACAG 22879
QY 2881 AGGAAGACCCCTTCCTTAAAGAAAGAAATGTGATTTTATCTTATGACGTACAGT 2940
Db 22880 AGGAAGACCCCTTCCTTAAAGAAAGAAATGTGATTTTATCTTATGACGTACAGT 22939
QY 2941 CATTAGTCATTAAGTTGAGTTGAGAGAAATATATGATCAGAGAAATTTATATCACT 3000
Db 22940 CATTAGTCATTAAGTTGAGTTGAGAGAAATATATGATCAGAGAAATTTATATCACT 22999
QY 3001 GTGTGCTGTAGATATATGAAAGAAATTAAGAGCTAGAGTCAAGGATTTCCATTAGTGA 3060
Db 23000 GTGTGCTGTAGATATATGAAAGAAATTAAGAGCTAGAGTCAAGGATTTCCATTAGTGA 23059
QY 3061 TTTGTTGTTTGTGTTTGAAGACAGAGTCTCTTTTGTTAACCAAGGCTAGAGTCAAGTGA 3120
Db 23060 TTTGTTGTTTGTGTTTGAAGACAGAGTCTCTTTTGTTAACCAAGGCTAGAGTCAAGTGA 23119
QY 3121 CAGTCAATGCTCAACGACGCTCAAACTCCAGAGCTCAAAATTTATCTTCCAGCTCGGCT 3180
Db 23120 CAGTCAATGCTCAACGACGCTCAAACTCCAGAGCTCAAAATTTATCTTCCAGCTCGGCT 23179
QY 3181 CCCAAAGTGTGGAATTTACAGGTGTGAGCCAAAGGTTTATTTGATGTGTCTGGCTAGT 3240
Db 23180 CCCAAAGTGTGGAATTTACAGGTGTGAGCCAAAGGTTTATTTGATGTGTCTGGCTAGT 23239
QY 3241 GCCTCTCAAACTTCAGTGAAGACAGAGTGAACGGGAACCTGACTCAACAGCTCGGGTT 3300
Db 23240 GCCTCTCAAACTTCAGTGAAGACAGAGTGAACGGGAACCTGACTCAACAGCTCGGGTT 23299
QY 3301 TTAAGCTGAAGCTCTGCTCATTTCAATGATCAAGTCAAGTGAAGTGTGCTGTCAAGA 3360
Db 23300 TTAAGCTGAAGCTCTGCTCATTTCAATGATCAAGTGAAGTGTGCTGTGTCAAGA 23359
QY 3361 GCCAAGCACTGAGAGCAAGAGATCTTGTTAGCAATTTAGTATCAAGTGTGATTTATGCT 3420
Db 23360 GCCAAGCACTGAGAGCAAGAGATCTTGTTAGCAATTTAGTATCAAGTGTGATTTATGCT 23419

QY 3421 AGTGACATTAAGATGAGAGGAATGTGAAATCAATTAACAAAGAGTTCACCTCTTG 3480
Db 23420 AGTGACATTAAGATGAGAGGAATGTGAAATCAATTAACAAAGAGTTCACCTCTTG 23479
QY 3481 GTTAATGTGACATGAGAGGAGGAAGATGTGGGCTCAACCTTAACCTGTTTGTGTTTGA 3540
Db 23480 GTTAATGTGACATGAGAGGAGGAAGATGTGGGCTCAACCTTAACCTGTTTGTGTTTGA 23539
QY 3541 GACGAGGAATTTGAGCTCTATTTAACAGAAATTAGAGAGAAAGTTGTTTGAAGAGAA 3600
Db 23540 GACGAGGAATTTGAGCTCTATTTAACAGAAATTAGAGAGAAAGTTGTTTGAAGAGAA 23599
QY 3601 GAGAGTCTCTGTTTCAAGAGTGTGAGGCTCCAGGTGACAGCATCTTCAAGAGGAAT 3660
Db 23600 GAGAGTCTCTGTTTCAAGAGTGTGAGGCTCCAGGTGACAGCATCTTCAAGAGGAAT 23659
QY 3661 GAGCAGTAGGCAACCTTAAAGAAATCTGTGCTCAAGAGGAGCTGTGAGCTCGAGTGT 3720
Db 23660 GAGCAGTAGGCAACCTTAAAGAAATCTGTGCTCAAGAGGAGCTGTGAGCTCGAGTGT 23719
QY 3721 AGATCTGAGGGTCAATCAGCATAGAGTTTGAAGACAAAGATAGGCAACCAAAAGAGC 3780
Db 23720 AGATCTGAGGGTCAATCAGCATAGAGTTTGAAGACAAAGATAGGCAACCAAAAGAGC 23779
QY 3781 AATATACAAAGAGAGAGAGATGATGATGAGACTTTTGCTTTTATGAGTGAAGAGGA 3840
Db 23780 AATATACAAAGAGAGAGAGATGATGATGAGACTTTTGCTTTTATGAGTGAAGAGGA 23839
QY 3841 ACAGGAAATGAAGATGAAGAGGAGAGAGCTGTGAGAAATGAGAGCATCTTGAAGAAAAA 3900
Db 23840 ACAGGAAATGAAGAGATGAAGAGGAGAGAGCTGTGAGAAATGAGAGCATCTTGAAGAAAAA 23899
QY 3901 ATACACACTGTATGAGAAATCAAGGAGAAAGAAATTTCAAGAGAGGGTATGTGACA 3960
Db 23900 ATACACACTGTATGAGAAATCAAGGAGAAAGAAATTTCAAGAGAGGGTATGTGACA 23959
QY 3961 GTATTTACAAAGATCAAGAAATNAGCTTAAAGCATACCTTGAACGANTGACCTTGTGG 4020
Db 23960 GTATTTACAAAGATCAAGAAATNAGCTTAAAGCATACCTTGAACGANTGACCTTGTGG 24019
QY 4021 ATTTGTAGGGGACACACTAATTAATTAAGAAATTTATGTGGGTATATGAGAGGCAAG 4080
Db 24020 ATTTGTAGGGGACACACTAATTAATTAAGAAATTTATGTGGGTATATGAGAGGCAAG 24079
QY 4081 GAAAGGTTATTCAAATCAAGAGAGTGGAGTGAAGTGTCTCAAGGTGAAGGCA 4140
Db 24080 GAAAGGTTATTCAAATCAAGAGAGTGGAGTGAAGTGTCTCAAGGTGAAGGCA 24139
QY 4141 TCAATGAAATGTGGAAAGGGGACAGAGCATCCATGCCATCCAGCAAGCCACTCA 4200
Db 24140 TCAATGAAATGTGGAAAGGGGACAGAGCATCCATGCCATCCAGCAAGCCACTCA 24199
QY 4201 GAAAGCTCCATGAGAGTTCAAGTATCAAGAGGCTCTGTACCTTAATCTTCTGGGTTT 4260
Db 24200 GAAAGCTCCATGAGAGTTCAAGTATCAAGAGGCTCTGTACCTTAATCTTCTGGGTTT 24259
QY 4261 TGCATAGGCTTATTTGTGAGGATGATTTATTTAACTATTTGGCCACTGTGATTAATT 4320
Db 24260 TGCATAGGCTTATTTGTGAGGATGATTTATTTAACTATTTGGCCACTGTGATTAATT 24319
QY 4321 AACCTTCAACCCCTGCCCTCCCTAATCATGACCTGTGCTTCCAGAGCAAGCCCTCA 4380
Db 24320 AACCTTCAACCCCTGCCCTCCCTAATCATGACCTGTGCTTCCAGAGCAAGCCCTCA 24379
QY 4381 TCCATAGCTACCAATGTGTGTCAGCTATCAAGTCACTTACAAAGAAAGCATCACTTG 4440
Db 24380 TCCATAGCTACCAATGTGTGTCAGCTATCAAGTCACTTACAAAGAAAGCATCACTTG 24439
QY 4441 GAGATTTCAAGATTTTGAAGTTGGCTGTCAAGAAATTTAGTTGAAGATCAAAATATAT 4500
Db 24440 GAGATTTCAAGATTTTGAAGTTGGCTGTCAAGAAATTTAGTTGAAGATCAAAATATAT 24499

QY 4501 TTCACATATACAGTCTGCTATTTTATATACAGGCGCCATTAATGCTTTTAAACAAAG 4560
DB 24500 TTCACATATACAGTCTGCTATTTTATATACAGGCGCCATTAATGCTTTTAAACAAAG 24559
QY 4561 AGGTGATAAATCAGATTTTCTTTTATATAAGGCTTACATGATGACAGTGTGTAATAG 4620
DB 24560 AGGTGATAAATCAGATTTTCTTTTATATAAGGCTTACATGATGACAGTGTGTAATAG 24619
QY 4621 ATTGGATGAGGCAATCTTTTATTTTGAATGTTATATTTCCCTGACCTTACTTCTC 4680
DB 24620 ATTGGATGAGGCAATCTTTTATTTTGAATGTTATATTTCCCTGACCTTACTTCTC 24679
QY 4681 CTGTTTTCTTCTACTCTCTCTCCCTACTACACAGAAACCTTCTCTCCCTACTCTCAT 4740
DB 24680 CTGTTTTCTTCTACTCTCTCTCCCTACTACACAGAAACCTTCTCTCCCTACTCTCAT 24739
QY 4741 TCCCTGAATGCTGTGCTGTTAAGGTTCCAGGCTTGAACAGTGAAGGCTTAATCAGAACAC 4800
DB 24740 TCCCTGAATGCTGTGCTGTTAAGGTTCCAGGCTTGAACAGTGAAGGCTTAATCAGAACAC 24799
QY 4801 AGTGTACAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4860
DB 24800 AGTGTACAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24859
QY 4861 GCGGGAACGACGCGGAAAAACAGAGTCAAGGCTATTTTCTGATCTGATATTTAA 4920
DB 24860 GCGGGAACGACGCGGAAAAACAGAGTCAAGGCTATTTTCTGATCTGATATTTAA 24919
QY 4921 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4980
DB 24920 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24979
QY 4981 ATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5040
DB 24980 ATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25039
QY 5041 GCTCCGACTGTGTGCGCTTAAAGTAAAGGCAATCTGATGATGATGATGATGATGATGATGAT 5100
DB 25040 GCTCCGACTGTGTGCGCTTAAAGTAAAGGCAATCTGATGATGATGATGATGATGATGATGAT 25099
QY 5101 AATCATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5160
DB 25100 AATCATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25159
QY 5161 ATGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5220
DB 25160 ATGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25219
QY 5221 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 5280
DB 25220 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 25279
QY 5281 ACGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5340
DB 25280 ACGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25339
QY 5341 ACCTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5400
DB 25340 ACCTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25399
QY 5401 CAGGACCCACCACTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5460
DB 25400 CAGGACCCACCACTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25459
QY 5461 ATGTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5520
DB 25460 ATGTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25519
QY 5521 AAAGTCTGGGATTTACAGGATGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5580
DB 25520 AAAGTCTGGGATTTACAGGATGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25579
QY 5581 ATGGGGTCTTGTCTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5640

DB 25580 ATGGGGTCTTGTCTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25639
QY 5641 GCTTGGCTCTCAAGATTTCTGGATTTACATGATGATGATGATGATGATGATGATGATGATGAT 5700
DB 25640 GCTTGGCTCTCAAGATTTCTGGATTTACATGATGATGATGATGATGATGATGATGATGATGAT 25699
QY 5701 CAATATTTAGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5760
DB 25700 CAATATTTAGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25759
QY 5761 AAAAATTAGAGATTTGGAACCAAAAGAGAGGCTCCAGAGAGGAGGAGGAGGAGGAGGAGGAG 5820
DB 25760 AAAAATTAGAGATTTGGAACCAAAAGAGAGGCTCCAGAGAGGAGGAGGAGGAGGAGGAGGAG 25819
QY 5821 ATGATCCCAATGCTGCAAGATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 5880
DB 25820 ATGATCCCAATGCTGCAAGATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 25879
QY 5881 GTCATGGAACCAATGCTGCAAGATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 5940
DB 25880 GTCATGGAACCAATGCTGCAAGATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 25939
QY 5941 TATAGGGGATGATTTTGTATTTTCAAGAGACAAACAGTTCATACATGCAAGATCTA 6000
DB 25940 TATAGGGGATGATTTTGTATTTTCAAGAGACAAACAGTTCATACATGCAAGATCTA 25999
QY 6001 GTGTGTACACAGGAGT 6060
DB 26000 GTGTGTACACAGGAGT 26059
QY 6061 GGT 6120
DB 26060 GGT 26119
QY 6121 TGCCCAATCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6180
DB 26120 TGCCCAATCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 26179
QY 6181 GAAGAGTATGAGAGATGTGG 6201
DB 26180 GAAGAGTATGAGAGATGTGG 26200

RESULT 3
US-10-027-632-4668
; Sequence 4668, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: US 60/135,720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4668
; LENGTH: 454

```

TYPE: DNA
ORGANISM: Human
US-10-027-632-4668

Query Match
Query Similarity 9.3%, Score 452, DB 13, Length 454,
Best Local Similarity 99.6%, Pred. No. 1,4e-94,
Matches 452, Conservative 1, Mismatches 1, Indels 0, Gaps 0,

QY 1944 CTTAAACACATCGCATTTTCAACAGCTCGATATAGAGAGCATGTTTACAGGAGAGAAAT 2003
DB 1 CTTAAACACATCGCATTTTCAACAGCTCTGATAGAGAGCATGTTTACAGGAGAGAAAT 60
QY 2004 GTTTTCAGCAAAAGGTACAGAGTACAGGAAATAGAGATATGTTTCAAGAAAGAACCCAG 2063
DB 61 GTTTTCAGCAAAAGGTACAGAGTACAGGAAATAGAGATATGTTTCAAGAAAGAACCCAG 120
QY 2064 AGCATGTTTGTATGAGGTTTGAAGGAAACACAGTGTGTTTGAACCTCCAGTTCATTAG 2123
DB 121 AGCATGTTTGTATGAGGTTTGAAGGAAACACAGTGTGTTTGAACCTCCAGTTCATTAG 180
QY 2124 TGCCTTATGAATCAATATGCTGTGTTACCACTGCACTTTTAAAAATGAAATTAATGGA 2183
DB 121 TGCCTTATGAATCAATATGCTGTGTTACCACTGCACTTTTAAAAATGAAATTAATGGA 240
QY 2184 TGAGAAAGATAGAAATATTTAGCATGTCATTTGTTGAAAGGCAAGTATATTTC 2243
DB 241 TGAGAAAGATAGAAATATTTAGCATGTCATTTGTTGAAAGGCAAGTATATTTC 300
QY 2244 TGCAACTTTTGTCCCAATGTGACTGTACTTATTTTATGATATGATGATGCAATACAG 2303
DB 301 TGCAACTTTTGTCCCAATGTGACTGTACTTATTTTATGATATGATGATGCAATACAG 360
QY 2304 ATACATATATATTTCTTACTGTAGACTGCACTGCAAAAATCTTTTAAAGCATGCGCTGCT 2363
DB 361 ATACATATATATTTCTTACTGTAGACTGCACTGCAAAAATCTTTTAAAGCATGCGCTGCT 420
QY 2364 CTTACTTCCTTATTTTTCAGAGAGAAATCCCAAG 2397
DB 421 CTTACTTCCTTATTTTTCAGAGAGAAATCCCAAG 454

RESULT 4
US-10-027-632-4668
Sequence 4668, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 4668
LENGTH: 454
TYPE: DNA
ORGANISM: Human
US-10-027-632-4668

```

Query Match	7.3%;	Score 452;	DB 15;	Length 454;
Best Local Similarity	99.6%;	Pred. No. 1.4e-94;		
Matches 452;	Conservative 1;	Mismatch 1;	Indels 0;	Gaps 0;
QY 1944	CTTAAACACACTCATTTCAACAGCTAGAAATAGAGAGCATGTTACAGGAGAGAAAT			2003
Db 1	CTTAAACACACTCATTTCAACAGCTAGAAATAGAGAGCATGTTACAGGAGAGAAAT			60
QY 2004	GTTTTCAGCAAGGTACAGGTAGGAGAAATAGAGGATGTTCAAGGAAGAGCCCG			2063
Db 61	GTTTTCAGCAAGGTACAGGTAGGAGAAATAGAGGATGTTCAAGGAAGAGCCCG			120
QY 2064	AGTCATGTTGTAGGGTTAGAGAAACACAGTGTTCGATCTCCAGGTTCAATTAG			2123
Db 121	AGTCATGTTGTAGGGTTAGAGAAACACAGTGTTCGATCTCCAGGTTCAATTAG			180
QY 2124	TGCGTTATGAAATCAATATAGTGGTTAGCAACCTGCATTTTAAATAATGAAATTAATG			2183
Db 181	TGCGTTATGAAATCAATATAGTGGTTAGCAACCTGCATTTTAAATAATGAAATTAATG			240
QY 2184	TGAGAAAGAAATAGAAATATATAGCATGCACTTACATTTGAAAGACAGATTAATTTTC			2243
Db 241	TGAGAAAGAAATAGAAATATATAGCATGCACTTACATTTGAAAGACAGATTAATTTTC			300
QY 2244	TGCAACTTTTGCCTGCATTTGAACTGTAATTAATTTTATATGATGATGGAATTAACG			2303
Db 301	TGCAACTTTTGCCTGCATTTGAACTGTAATTAATTTTATATGATGATGGAATTAACG			360
QY 2304	ATACATATATATTTCTTACTGTAGACGTGCATCAAAAATCTTTTAAAGCATGCGCTGCT			2363
Db 361	ATACATATATATTTCTTACTGTAGACGTGCATCAAAAATCTTTTAAAGCATGCGCTGCT			420
QY 2364	CTAACTTCCTATTTTGCAGAGGAGAAATCCCAAG 2397			
Db 421	CTAACTTCCTATTTTGCAGAGGAGAAATCCCAAG 454			

```

RESULT 5
US-10-322-281-751
/ Sequence 751, Application US/10322281
/ Publication No. US20040126762A1
/ GENERAL INFORMATION:
/ APPLICANT: David W. Morris
/ APPLICANT: Marc S. Malandro
/ TITLE OF INVENTION: Novel Compositions and Methods in Cancer
/ FILE REFERENCE: 52945200100
/ CURRENT APPLICATION NUMBER: US/10/322,281
/ CURRENT FILING DATE: 2002-12-17
/ NUMBER OF SEQ ID NOS: 866
/ SOFTWARE: PaetSeq for Windows Version 4.0
/ SEQ ID NO 751
/ LENGTH: 2403
/ TYPE: DNA
/ ORGANISM: Homo sapiens

Query Match          6.5%; Score 402.6; DB 17; Length 2403;
Best Local Similarity 92.5%; Pred. No. 1,1e-82;
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

16 CAATTATGAAGAAGCCAAATATATACGCCCTTAATAATGATGTGGACCTTAATATACCGAAG 75
|||||
465 CAATTATGAAGAAACCGAAATATACACCTTAAATATGAATGTGGACCTTAATATCCGAG 524
|||||
76 AACTCCCCCTTTGTAGATTGTTGAACAAAATTAATATGATAGTAGATTAAATGTTCTAATG 135
|||||
525 CACCCCCCTTTGTAGATTGTTGAACAAAATTAATATGATAGTAGATTAAATGTTCTAATG 584
|||||
136 GAATGTGAACCCAGACGCAATATAGGGCTAGCAAAATGGAGAAATCATATATCATCA 195
|||||
585 GAGTGTGACCCAGACGCAATATAGGTCTAGCAAAATGGCAGAAATCATATATAGCATCA 644
|||||

```

QY	196	AAGTATCTCTCAAGAGCTTCAGCGCTATATATGCTAAAGAAATGTGAACGCCCTC	255
Db	645	AAGTTGTCTCGACAGACTTCGGCGCTATATATGTCTAAAGAAATATGAATCTCTTC	704
QY	256	AGCCATCTGAAGAGACAGTGTATCAGCAATTGATCAAAAAAGAAAACCAAGCGCCCTCC	315
Db	705	AGCGCCCGGAAGAGACAGTGTACACCAATTATCAAAAAAGAAAACCAAGCGCCCTCC	764
QY	316	CTTCCCCCATCTGATGTAAAGACAGCTTCATTTCCATAGTAAATTTCTAAATAC	375
Db	765	CTTCCCCCATTCGATTTATCAAGCTTCATTTCCATAGTAAATTTCTAAATAC	824
QY	376	AGCTTGTAGACTCAAGTACTGGAAGAAAGACCTCCATTCAAAGAAATTTATCTTAAG	435
Db	825	GCTTGTAGACTCAAGTACCGGAAGAAAGACCTCCATTCAAAGAAATTTATCTTAAG	884
QY	436	ATACGTAAATGATATCAAA-TTTTGTGACATTGGAAATATTAAGTGT	483
Db	885	ATACGTAAATGATATCAAAATTTTGTGACATTGGAAATATTAAGTGT	933

```

RESULT 6
US-09-814-353-20512
: Sequence 20512, Application US/09814353
: Publication No. US20030165831A1
: GENERAL INFORMATION:
: APPLICANT: Lee, John
: APPLICANT: Thompson, Pamela
: APPLICANT: Lillie, James
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
: TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
: TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
: FILE REFERENCES: MRI-006B
: CURRENT APPLICATION NUMBER: US/09/814,353
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: US 60/191,031
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: US 60/207,124
: PRIOR FILING DATE: 2000-05-25
: PRIOR APPLICATION NUMBER: US 60/211,940
: PRIOR FILING DATE: 2000-06-15
: PRIOR APPLICATION NUMBER: US 60/216,820
: PRIOR FILING DATE: 2000-07-07
: PRIOR APPLICATION NUMBER: US 60/220,661
: PRIOR FILING DATE: 2000-07-25
: PRIOR APPLICATION NUMBER: US 60/257,672
: PRIOR FILING DATE: 2000-12-21
: NUMBER OF SEQ ID NOS: 22037
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 20512
: LENGTH: 3802
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-814-353-20512

```

Query Match	6.5%	Score 402.61	DB 10	Length 3802
Best Local Similarity	92.5%	Pred. 1.4e-82		
Matches	434	Conservative	0	Mismatches 34; Indels 1; Gaps 1;
QY	16	CAATTATGAAAGCCAAATATACAGCCTTAAATATGATGTGGACCTTAATATCCAGAAG	75	
Db	438	CAATTATGAAAAACCGAATATACGCTTAAATATGAAATGTGGACCTTAATATCCAGAAG	497	
QY	76	AACCTCCCTTGTAGATTGTGAACAAAATTAATATAGTAGTAGTAAATAGTCTTAATG	135	
Db	498	CACCCCTTGTGAAGATTGTAAACAAAATTAATATGAATGAGTAATAGTCTTAATG	557	
QY	136	GAATGTGAACCCAGAAGCCATATACAGGCTAGCAAAATGCGACAAATCATATATCATCA	195	
Db	558	GAGTGTGTGACCCAGAAGCCATATACAGCTAGCAAAATGCGACAAATCATATAGCATCA	617	
QY	196	AAGTTATCTTCAAGAAGCTTCAGGCGCTTAATGATGCTTAAGAAATGTGAAGCGCCTC	255	

Db	618	AAAGTTGTCCTGCACAAGAGCTTGCGGCGCTTAATGATGTCCTAAGAAAAATATGAACACTCCCTC	677
Oy	256	AGCCATCTGAAGGACAGCTGTTACAGCAATTTGATCAAAAAAGAAAAACAAGGCCCTTTCCC	315
Db	678	AGCGCGCCGGAAGGACAGGTATACAGCAATTTATCAAAAAAGAAAAACAAGGCCCTTTCCC	737
Oy	316	CTTCCCCCATCTTGATGTAAGCAGCTCTCAATTTTCCATAGTAGTAATTTTCTAGATAC	375
Db	738	CTTCCCCCATTTGATTTAATCACTCTTCATTTTCCACAGTAGTAATTTTCTAGATAC	797
Oy	376	AGCTTTGAGACTCAAAAGTACTGGAAGAAAGCTCCCATTCAGAGAAATTTATCTTAAG	435
Db	798	GCTCTGTAGACCTCAAAAGTACCGGAAGAAAGAGCTCCCATTCAGAGAAATTTATCTTAAG	857
Oy	436	ATTACTGTAAATGATATCAATA-TTTTGTGATCAATTGGAATATATAAATTGT	483
Db	858	ATTACTGTAAATGATATCAATAATTTTGTGTCATTTTGAAATATATAAATTGT	906

```

RESULT 7
US-10-087-192-904/C
: Sequence 904, Application US//10087192
: Publication NO. US20020182586A1
: GENERAL INFORMATION:
: APPLICANT: Morris, David W.
: APPLICANT: Engelhard, Eric K.
: TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
: TITLE OF INVENTION: CANCER
: FILE REFERENCE: 529452000122
: CURRENT APPLICATION NUMBER: US//10/087,192
: CURRENT FILING DATE: 2002-03-01
: PRIOR APPLICATION NUMBER: US 09/747,377
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: US 09/798,586
: PRIOR FILING DATE: 2001-03-02
: NUMBER OF SEQ ID NOS: 2059
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 904
: LENGTH: 75252
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-087-192-904

```

Query Match	6.0%	Score 375;	DB 13;	Length 75552;
Best Local Similarity	57.4%	Pred. No.1.8e-75;		
Matches 873;	Conservative	0;	Mismatch 625;	Indels 23; Gaps 10;
QY 483	TTAGCCCTGCGCGCGGTGGCTCAGCCCTGTATCCACGCCCTTTGGAGGCGCAGAGTGAGCA	542		
Db 31345	TTGGTCGGGCGCGGTGGCTCAGCCCTGTATCCACGACTTTGGAGGCGCGAGGCCGGC	31288		
QY 543	GATCATGAGGTGAGGATTTGAGACCCAGCCTAGCCCAACATGTGTAACCCCGTCTTACT	602		
Db 31285	GATCAGAGGTGAGGAGATGGAACCATCTGTCAACACGGTGAACCCCGTCTTACT	31222		
QY 603	AAAAGATAC--AAAAATTTAGCCAGGTGTGTGTGGCGCACACTGTAAACCCAGTCTCGA	660		
Db 31225	AAAAATACAAAAAAAATTTAGCCAGGGGTGTGTGGCGCGCTGTATCCAGCTTACTCG	31166		
QY 661	GAGAGTGAGGCGAGAAATTGCTTTGAACCGAGAGGCGAAG--GTGAGGAGCGAAGAATGC	719		
Db 31165	GAGGCTGAGGCGAGAAATGCGCTGTAAACCCGGAGGCGGAGCTTGAAGTGAGCGAGATGC	31106		
QY 720	ACACCAATGACTGTAGCTGTAGTGACAGGGCAAGACTCAACTCAAAAAAAAAAAAAAAA	779		
Db 31105	GCACTACTGACTCTCGAGCTGGATGACAGATGAGACTCCGTCTCGGAAAAAAAAAAAAAAA	31046		
QY 780	AAAGAAATATGTAAGTTGTGC---TATTAACAATTAATAGCGAGTGAGAACAAAGTGC	835		
Db 31045	AAATACAAAATCTTAAGCTGGGCGCAGTGCGAGACGCTTTGTAATCCAGCTTATTAAGAGGC	30986		
QY 836	TAAAGCTATGACCATGTGTACTTAGAATACTGTGGAAACACTTAATAAGGGAACCTTAAC	895		

Db 30985 TGAAGCAGGAGATGTTTGAACCTGGAGGCAAAAGTTGCAAG-TAAGCCAAAGATCACGCC 30927
QY 896 CCAAGTCTGGAAAGTATTTTGGAAAAGAAATGTTTGGAGCAAAAGGTTTAAAGAGATG 955
Db 30926 ACTGCACTCCAGCTGGGTGA CAGAGGAGACTGTGCTCAAAAACAATAATTAATAATAA 30867
QY 956 AAAAAAAAAATTAATAATCAAGTTTACCTGTGTGAGAAATGGATAGGAGCTTAACCTAGA 1015
Db 30866 ATATAATTAATAATTAATAATTAATAATAATAATAATAATAATAATAATAATAATAATA 30807
QY 1016 GAAATCAATAGGAATGTTTCATGTATGTTTAAGAACCTGGTAAGGATGAAGACATTA 1075
Db 30806 GTGGCTGAGACATAGCCCTTGTGTGGATCTTAAGGTTTAAATCACTGGGTGCTTACT 30747
QY 1076 CATTAATGCAACATCGCGGACTTTTATTTATGTAATGCTTGGCAATTTAATAGAG 1135
Db 30746 GGGCTCTTGTGTGGGGAGGAGGAGCTCCAAG---AGGTGGCTCAAGGAGATAGAA 30690
QY 1136 GAGCAGGAATATGAGATGTTGATTAAGTCAAGATGTAAGTTTGCACAACATGTGAAA 1195
Db 30689 TGAAGCAGCTGGAGGAGGAGTGGGCGCTGGGACTGGGTGACTCTTGGCTTGGCTGGG 30630
QY 1196 GGAAGAGACAGTATGAGCAAGAGATTATCAAGACAGAAAGTTAATGTGCTGG 1255
Db 30629 GAAAGAGAGAGAGGCGCATCACTAGCAACAGAGCAAGAAAGAGAGAGTGTGGAGCC 30570
QY 1256 CCAATGGCATCTAGTCTTGAAGTCTTAATCTGAGGAGAGAGTGAAGATTAAGAGCTTGTG 1315
Db 30569 CCAAGCCCTCAAGATCTGTGTACACTCTCCCAATGAGAGAGAGGAGGAGGAGGAGGAG 30515
QY 1316 ATAGTTATGAAGAGAGTGAAGAGGCTCAAGAGCTTAAGGTTGATTAATGAAGAAT 1375
Db 30514 CATGAGCTGAAGATGTGCAATTTTCTCAACGTGTACCAACCGAATGTGTGCAACAACCA 30455
QY 1376 GATTGAGAAAGATTAATCTGTGAGAGAGTGAAGTTCAGGCTTGAAGTCACTCTCATATA 1435
Db 30454 AGTGTGGGGCTGTGTAGCCAGAGAGACTGACAGCAT---CGATTATATGACTGAAGC 30399
QY 1436 CCAAGCATCTGTCTAATGCTTCAAAAGACATGATCCTGCTTCAAGGAGCTTAAGAGCCA 1495
Db 30398 ACAGATGGAGTGAAGCCCAAAAGACATGATTAAGTGAAGCTGAAGGAGTTCACAGGCA 30339
QY 1496 AAAACAAGAGATA-AGAAATACACACCAATACTATTATAGACACTGTGTAGAATATCA 1554
Db 30338 GGGAAATACACAGGAGAGAGAGAAAGACACACTTCTCAGACAGAAATGGAAAGAGG 30279
QY 1555 AGAAAGAAATACGATCTAGTACTGTAGATGTGAACGGCATCAAGATATCTTCTAGTTT 1614
Db 30278 GGAAGAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 30219
QY 1615 CAAGAATTTCAATGCGCGCGCGCGCGCGCGCTCAAGCTTGTATCCACAGACTTTGGGA 1674
Db 30218 CAATAAATAAATAATGCGCGCGCGCGCGCGCTCAAGCTTGTATCCACAGACTTTGGGA 30159
QY 1675 GGGCGAGGCGGGTGAATCAACAAGTCAAGAGTCAAGACCATCTGGTTTAAACGGTGA 1734
Db 30158 GGGCGAGGCGGGTGAATCAACAAGTCAAGAGTCAAGACCATCTGGTTTAAACGGTGA 30099
QY 1735 ACCCGCTCTTAC-AAAAATATATAATAATAATAATAATAATAATAATAATAATAATAATA 1793
Db 30098 ACCCGCTCTTACAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 30039
QY 1794 CCAAGTACTCAGAGGCTGAAGGAGAGAGATGGCGTGAACCCGGAGAGTGAAGG-TTTGC 1852
Db 30038 CCAAGTACTCAGAGGCTGAAGGAGAGAGATGGCGTGAACCCGGAGAGGAGGAGCTTTCA 29979
QY 1853 GTGAGCGAGATGCGCGCACTGCGCTCAGGCTGCGGAGCAGAGTGAAGTGGGTGCA 1912
Db 29978 GTGAGCGAGATGCGCGCACTGCGCTCAGGCTGCGGAGCAGAGTGAAGTGGGTGCA 29919
QY 1913 AAAAAAAAAAAAAAAAAAGTTTCAATCTTAAACACATCTTCAACAGTCTAG 1972
Db 29918 AAAAAAAAAAAAAAAAAAGATCTGAATATATTAATTAATTTGTGACAAACATTA 29859

QY 1973 AATAGAGAGCATGTTACAGG 1993
Db 29858 AGGTGAGACTACATAGTACGGG 29838
RESULT 8
US-10-087-192-274/c
; Sequence 274, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 73145
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(73145)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-274
Query Match 6.0%; Score 373.6; DB 13; Length 73145;
Best Local Similarity 51.7%; Pred. No. 3.8e-75;
Matches 1278; Conservative 0; Mismatches 1104; Indels 92; Gaps 15;
QY 449 TACTAATTTTGTGACATTTGGAATATATTAATGTTGTTAGCTGGCGGCTGCTACGCTT 508
Db 43054 TGCATATATTTCAATATATATATGAAGAGAAACACAGGCTGGGCAAGTGTCTCACACT 42995
QY 509 GTAATCCCGCCCTTGGGAGGCGAGAGTGGGAGAT--CATGAGTCAAGAGTTGAGA 566
Db 42994 GGAATCCCAACTTTGAGAGGCGCAACATGGAGATGCTTGAAGCCAGAGTTCAAGA 42935
QY 567 CCAAGCTAGCAACATGTGTAACCCCGTCTTACTTAAGATACAAAATTAAGCCAGGT 626
Db 42934 CCAAGCTGGGCAACACAGTGAACCCCATCTTACAAAATTAAGAAACAATTAAGCAGAT 42875
QY 627 GTGTGGCGCACACTGTATACCCAGCTGTGAGAGATGAGGAGAGAAATGCTTGA 686
Db 42874 GATGTGTGTCATGCTGTATGTTCCAGCTACTTGGAGGCTGAGGTGGGAGATCTTTGG 42815
QY 687 ACCCAGAGGAGAGAGG-TGACAGAGCAAAAGATCAACCAATGCACTGTAAGCTGATGA 745
Db 42814 GCCCAGAGGAGGAGGCTGCAAGTGAAGCTATGATCAACCACTGCACTCCAGCTT----- 42761
QY 746 CAGGCGAAGACTCAACTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAATATATGTTGTATTA 805
Db 42760 -----GTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAGAAAGAAACACATGGT 42709
QY 806 CAAATTAATAGCAGTGAAGAGCAAAAGTCTTAAGCTTATGCAATGTAATAGAAATA 865
Db 42708 GTTTCACAAATTAAGAGAAAGAAAGAAAGTAATCAGCATGAAAGATTAACAAATTAAGACA 42649
QY 866 CTGTGGGAAACATATATAAGGAGCACTTAACCAAGTCTGGAAGTAAAGTTTGAAGAGGA 925
Db 42648 AATTTGAGTGTGCTGAAGATTAAGAAAAATACAACT-----TAA 42611
QY 926 ATGTTTGAAGCAAAAGGTTTAAAGAGAGTGAAGAAAAAAATTAATAATCAAGTTAGCTG 985
Db 42610 TTTCCTGTGTAAAGAGCTTGGCACTTCCAGAAAGAAAAATGAATTCAGCAATTAAGCT 42551

QY 986 TGTGGAATGGGANTGGAGCTAATGAGAAATCAATAGGAATGTTTCATGTATGT 1045
DB 42550 GTGGGCAAGAGAGTCAATTAACAAATATGGAAGCTAAAAAGAGTTTCACTAACTT 42491
QY 1046 TTAGGACCTGTGAAGGTGAAGACATTAATCATTAATCTGACATGCGGGAATTTTTT 1105
DB 42490 CAAATGTTGAAAGGCTGGATTCATATGTAATATATGCAAAAAATTAATTTAGCA 42431
QY 1106 TTAATGTAATGCTTGGCAATTTAAATAGAGAGCAGAGATGTGAACAGTTGATGAT 1165
DB 42430 AATGTAATGATGTATTAAGAGAGATATCAGATATAGAGGCAATGGGTGGCAGACAT 42371
QY 1166 CAGAGTTGAATGTTCTGCCAGACATGTGAAAGAGAGCAGGTAGGCAAGATTGAA 1225
DB 42370 ATTACATGGAATCTAATGCTAATGAACAGAGCAAAACATATATCTCAATATCTGAG 42311
QY 1226 AATTTATCAAGACAGAAATGTAATGTGCTGGCCAGTGGCATCTAGTCTGATCTAATCTGA 1285
DB 42310 AAAAAAGAGTGAATGATTAATGATTAATGATTAATTAATTAATTAATTAATTAATTA 42251
QY 1286 GGGAGAGAGTGAAGATGAAGCAGCTTGTGATGTTATGAGAGAGTGAAGGCTTCAAG 1345
DB 42250 ACATTAAGAAATTCAGACACATTAAGAGAGCTGAAGATCAAAACA-AGGTAAATTAATG 42192
QY 1346 GACCTACAGGTGTTGATTAATTAAGAAATGATTTGAGAAAGAAATACGTGAGAGATG 1405
DB 42191 GTTGTGTTGGTGGGATTAATTAATGAACAAACA---GATAATTTCTTCTTTTCAATAT 42135
QY 1406 AGATTTTCAGGCTTGAATGATCTCATACAGACATCTGTGTAATGCTTCAAGACA 1465
DB 42134 CATTTGATGTTCAAAAAAGTCAATCATATATAGTCTCAATCAATCAATCAATCAAT 42075
QY 1466 TGTATCCCTGCTCAAGAGACTTACAGCCAAAAAAGAGATGAAGATACA-CACCAAT 1524
DB 42074 CAAAAGCAAAATGATTTCAAGGCTCATTTCTCCCTACGCAAAAAAAGAGATTAAGAAC 42015
QY 1525 ACTATTAATAGGACACTTGTGTAGAAATATCAAGAAAGAAATACATCTAGTCTGATG 1584
DB 42014 AAACAGAAATAGAAATAGTATATTTTAAATATATCTTAAGAAAAATTTGTCTGA 41955
QY 1585 TGCAGGCAATCAAAATATCTTCTAGTTTCAAGAAATTTCAATGCGCGCGCGGCTG 1644
DB 41954 GAAAGAAATGCAAGAAACATATCTGTCTGTAAGAAATACATGCGCGCGCGCTG 41895
QY 1645 GCTCAGGCTGTATCCAGACCTTTGGAGGCGGAGCGGTGATCAAGAGTCAAGA 1704
DB 41894 GCTCAGGCTGTATCCAGACCTTTGGAGGCGGAGCGGAGTCAAGAGTCAAGA 41835
QY 1705 GATCAAGACATCTGTGTTAACAGGCTGAACCCGCTCTTACAAAAATATATTAAT 1764
DB 41834 GATCGAGACATCTGTGTTAACAGGCTGAACCCGCTCTTACAAAAATATATTAAT 41775
QY 1765 AGCAGAGCGTGTGGGGGCGCTGTAGTCCAGCTACTCAAGAGGCTAGAGGAGAGAA 1824
DB 41774 AGCGGCGCGTGTGGGGGCGCTGTAGTCCAGCTACTCAAGAGGCTAGAGGAGAGAA 41715
QY 1825 TGGCGTGAACCCGAGAGTGAAGTTTGC-CTGAGCGAGATCGCGCACTGCGCTC-CA 1882
DB 41714 TGGCGTGAACCCGAGAGGCGGAGCTTGCAATGAGCGAGACATGCACTGCACTCTCAC 41655
QY 1883 CTTGGGCGAGAGAGTGAAGTGGCTCTCAAAAAATTAATTAATTAATTAATTAATTA 1942
DB 41654 TCTGGGCTGACAGAGTGAAGTCTCGCTCTCAAAAAATTAATTAATTAATTAATTA 41595
QY 1943 TCTTAACACACATGCTATTTCAACAGTCTAATTAAGAGAGCATGTTACAGGAGAGAAA 2002
DB 41594 CTGACATATCAAAACCTATAGAAAGATGAATGAG-----GAAAG 41554
QY 2003 TGTATTCAGCAAGATGACAGAGTGAAGAAATAGAGATATGTTCAAGAGAGAGACCCCA 2062
DB 41553 TCTTGAAGAAACACATCAAGCTTAATATGCTTTATTTTCAAGAAAGATGAGAAA 41494

QY 2063 GAGTCAATGTTTGTAGGTTAGAGAAACAGAGTGTGATGCAATCTCAGGTTCCATTA 2122
DB 41493 AA-----TGATGAACCTAAGATTAATCAATCTTGAAGAAAGAA 41453
QY 2123 GTGCGTTATGAATCAATATGTTGTTAGCAACTGCAATTTTAAAAATGAATTAATG 2182
DB 41452 CAATTTTAAAGAAAGGAAAGGAGGAGAGTGAATGATTAAGCAAAAAA----- 41397
QY 2183 ATGAGAGAGAAATGAAAAATTAAGCATGATTAATGTTGAAGAGCAAGTATTTT 2242
DB 41396 CTGAGAAAGAGACAGGTGATGCAAGAGTGGCTTATTTCTTAAGACTCAAAACTGG 41337
QY 2243 CTGCAACTTTTCTCCAAATGTAATGTAATTAATTTTATGATGATGAATPACA 2302
DB 41336 ATGTGGGTTCAATTAACAGTGGCTCTCAGTTCTTTGATGAATGAATGATCTTA 41277
QY 2303 GATACATATATATTTCTTCTAGTACAGTCAAGTCAAAAAATCTTTAAAGCACTGG 2362
DB 41276 ACAGAAAGAAAGGAGAAAGAGAAACACACAAAGAGAGGAGCATTAAGCTTG 41217
QY 2363 TCTTAATCTTATTTTGAAGAGAGAAATCAAGATCTGAAGAGCAAAATTTGCTT 2422
DB 41216 AGCTTAAGGATATTAACACGTAAGAAATCTATGACAGATGTTTCAGTTAATGAGA 41157
QY 2423 GAGCTATTAAGACAGCTTATGCTATGCTTAAGATTTCTTATTAATTTCTTCC 2482
DB 41156 AATCTGAAGACATGACCAATTAACACAGATGAAGATTAACGTTAAACAAAGCTAG 41097
QY 2483 ACTAGGCAATGCTGCACTTCTAGTTCTGTGGCTGAATACAGAAATATTAATGATA 2542
DB 41096 AAGTCAAAATTAATCAATATGAGAGAGTAACTGAGAGAGAAATCCCCCAATTA 41037
QY 2543 CAGCATACACAGTCTGGGA--AAATATGAGTGAAGTGGCTGAGAGCTCATTTCTA 2600
DB 41036 GAGGTTGATGCTTATGATGAATTCATTCATACTTAAGAAATATGATATCTCAC 40977
QY 2601 AGAAATGAGACCTTAAGGAGGATATGTTGCTCACACTATTAATCCAGACTTTGGA 2660
DB 40976 ATCACAATTAATTTCTGCTGCTGCAAGTGCATGCTGTAATCCAGCACTTTGGA 40917
QY 2661 GGCAGAGTCAAGAGATGCTTGAACCCAGAGTTCAAGACTAGCATGAGCAATAGCA 2720
DB 40916 GGCAGAGGAGGATGATCTTGAAGCCAGAGTTTGAACAGCTCGGAAACACAGTG 40857
QY 2721 AGACTCATCTTACAAAAATTTAAATCACTGAGAGATGAGTATGCTGAT 2780
DB 40856 AAACCCATTTCTACAAAAATTAAGCAAGGCTGAGTGTGATGCTGAT 40797
QY 2781 CCCACTTACCTGGAAGCT-AGTGGTGAATGCTTG-ACACAGAGTTTGAAGCTAAG 2838
DB 40796 CCCAGTCTTACAGAGGCTGAGAGAGATCACTTGAACCAAGAGGTGAGGTTGCA 40737
QY 2839 GTGAGCATGATCAACACATGCACTGAGCTTGAAGTCAAGAGAGAGACCTGCTCTA 2898
DB 40736 GTGAGCATGATCAACACATGCACTGAGCTGAGTCAAGAGTCAAGTCACTGCTCTCA 40677
QY 2899 AAAAAAGAAAT 2912
DB 40676 AAAAAAATAAT 40663

RESULT 9
US-10-265-071-22
; Sequence 22, Application US/10265071
; Publication No. US20030147875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerlich, Stefan
; APPLICANT: Tomita, Megumi
; TITLE OF INVENTION: Sulfates and methods of use thereof
; FILE REFERENCE: US-10-230CIP
; CURRENT APPLICATION NUMBER: US/10/265,071
; CURRENT FILING DATE: 2003-01-22

PRIOR APPLICATION NUMBER: 60/258,577
PRIOR FILING DATE: 2000-12-27
PRIOR APPLICATION NUMBER: 60/267,831
PRIOR FILING DATE: 2001-09-02
PRIOR APPLICATION NUMBER: 10/025,966
PRIOR FILING DATE: 2001-12-21
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 268685
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_structure
LOCATION: (0)...(0)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (0)...(0)
OTHER INFORMATION: n = A,T,C or G
US-10-265-071-22

Query Match 5.9%; Score 365.2; DB 15; Length 268685;
Best Local Similarity 52.8%; Pred. No. 7e-73;
Matches 1289; Conservative 0; Mismatches 1043; Indels 108; Gaps 19;

QY 486 GCGTCGCGCGTGGCTGACCCCTGTAATCCAGCCCTTTGGAGGCGCAGGTGGCAGAT 545
DB 221964 GCGAGCGCGGTGGCTGCTGCTGCTGTAATCCAACTTTGGAGGCGCAGGTGGGAT 222023
QY 546 CATGAGGTGAGAGTTTGAAGCAGCAGCTAGCAATGAGAAACCCCGTCTGTAATAA 605
DB 222024 CATGAGGTGAGAGTTTGAAGCAGCAGCTAGCAATGAGAAACCCCGTCTGTAATAA 222083
QY 606 GATACAAAAAATTAGCCAGGTGTGTGGCGCAGACCTGTAAACCCAGCTGCTGAGAG 665
DB 222084 AATACAAAAAATTAGCTGAGTGTGTGGCGCAGACCTGTAAACCCAGCTGCTGAGAG 222143
QY 666 TGAAGCAGAGAGATTGCTTAACCCAGAGAGCAGAGG-TGCAGCAGCAAGATCACACC 724
DB 222144 TGAAGCAGAGAGATTGCTTAACCCAGAGAGCAGAGGTTGCAATGAGAAACCTGCGCC 222203
QY 725 AATGCACTGAGCCTGGATGATGAGAGGCAAGCTCAACTCAAAAAAAGAAAAAAGA 784
DB 222204 ACTGCACTGAGCCTGGAGCAGAGCAGAGCTCGCTCAAAAAAAGAAAAAAGAAAAA 222263
QY 785 AATATGTAAGTTGTGTATAACAATTAATAGGCAAGAGCAAAAGTCTAAAGCTTA 844
DB 222264 AATGTCCTGCGCATATATTTATCTTCTATGACATATTCATTAGAAAAATCATCTGAGA 222323
QY 845 TGACCATGTACTAGGAATATCTGTGGAAACATATTAAGGAACTTAACCCAGTCTTG 904
DB 222324 AATGTCCTGCGCATATATTTATCTTCTATGACATATTCATTAGAAAAATCATCTGAGA 222383
QY 905 G-AAGTAAGGTTTGGAAAGAAATGTTTGAAGCAAAAGGTTTAAAGAGTGAAGAAAAA 963
DB 222384 GCTACTGAGATTTAAAGCACTAATATTTGAATAGCTAGCAAGAAATTAACACAGTGC 222443
QY 964 AATTTAAATACAGTTTAAAGTGTGTGAGAAATGGATAGGAGCTTAAGAGAAATCAA 1023
DB 222444 CATTAATCTGTTTAAATCAGAAATTTAATGAGCAAAATTTATTAATAGCAGGG 222503
QY 1024 ATAGGATGTTTATGATGATGTTTAAAGGACCTGTGTAAGGTTGAAGACATTAATATCT 1083
DB 222504 GGAAGCTAGGAGAAAGGATTAACGCAATCTGGAAGAGATTAATCTGTGACTGCTG 222563
QY 1084 GCAACATCGGAGGACTTTTATGATGATGTTTAAAGGAGTGAAGTGAAGTGAAGG--AGCAG 1141
DB 222564 CCAAAATTAACATGATTTGTCTCAACCACTCTCTGTGATTTAGCCCTGGGCGACATG 222623
QY 1142 AGAATGTAACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
DB 222624 ACAAATTAACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 222683

QY 1201 AGACAGTAGGAGAGAGAGTTTGAAGATTTATCAGACAGAAATTAAAG-TGCTGGCCAG 1259
DB 222684 AATAAATATCCCAAGAGGTCAAAAGTATAGATCTAGTTTAAATATATGTAATCAAC 222743
QY 1260 TGGCATCTAGTCTGAGTCTAATCTGTGAGAGAGAAAGTGAAGTAAAGCAGCTGCTGATAG 1319
DB 222744 ACAAAGCATTTGATCTTCTGCTGGCAATTTTTTCTCCATCATGCTGTTTCTTT 222803
QY 1320 TTAATGAAGAGTGAAGGCTTCAAGACCTAAGGTGTTGAATTAATGAAGATGATTT 1379
DB 222804 CATGCTTGGTAAATTTCTAATTTTAAATAAATGATATCAGTTTATATCTAGATGAGA 222863
QY 1380 GGAAGAAAGATTAAGTGTGAGAGATGAGATTTTCAAGCTTGAAGTGTGATCTCATCCAG 1439
DB 222864 TAATTTTCA-----AGATGCTACTTGTCTTAATGCTTAAGATATTTT 222907
QY 1440 ACACGTGCTAAATGCTTCAAGACAGATGATCCCTGCTCAAGGAGCTTAAGCCAAAAA 1499
DB 222908 ATTTAGTATCTTCAGTAGCAATGTAACATGAGCTATTCAGATTAAGAAATTAACCTGC 222967
QY 1500 CAAGAGATTAAGAAATTAACACCAATATCTATTAAGACACTGTGTGAATATCAAGAAA 1559
DB 222968 CAATCATTAAGAGTGTGAGAGAAAGACT----CACTCTAAAGAGTAAATATATATATA 223023
QY 1560 GAAATTCAGTCTAGTCTGATGATGTGCAACGCAATCAAGATATCTTCTAGTTTCAAGA 1619
DB 223024 CATTTGCTGTCACTATGTTTATCTGAGGAGTTATATACAGAGAAATTTCAATATATA 223083
QY 1620 AGTTTGAATTCGCGCGCGGTGGCTGACCGCTGTAATCCAGCAGCTTTGGAGAGCGG 1679
DB 223084 AGTAACATA-CAGCGGCGATGTGTGCTCATGCTGTATTCACACACTTTGGAGGCGG 223142
QY 1680 AGCGGCGGTGATCAC--AAGTCAAGAGATCAAGACATCTGTGTTTAAACGCTGAAC 1737
DB 223143 AGCAGAGCGGATTTACCTGAGGTCAGAGCTTGAACCAACCTGACCAATGTGTAAC 223202
QY 1738 CGCTCTCAAAAAAATTAAAAAATTAGCAGGCGGTGTGGCGGCGCTGATGCCA 1797
DB 223203 CGTCTCTAATAAATGTAATAAATTAACAGGCGGTGTGATGGCGCTGTAATTTCA 223262
QY 1798 GCTACTCAGAGGCTGAGGAGAGAGATGCGTGAACCCGAGAGGTAGATTTGC-GTGA 1856
DB 223263 GCTACTCAGAGGCTGAGGAGAGAGATGCGTGAACCCGAGAGGTAGATTTGC-GTGA 223322
QY 1857 GCGAGATGCGCGCACTGCGCTTCAGCTGGG--GCAAGAGTGAAGCTGCTTCAAAA 1914
DB 223323 GCTGAGATCATGCGATTTGACCTTCAGCTGGGCGCAGAGCAAAATCTCGCTCAAAA 223382
QY 1915 AAAAAAAGAAAAAAGTTTCAAGATCTTAAACAGCTGATTTCAAGGTGAAA 1974
DB 223383 AAAAAAAGTAAAC-----TCATTCAGCGTAAAC 223413
QY 1975 TAGAGAGCATGTTCAGAGAGAGAAATGTTTTCAGAAAGTACAGAGTGAAGAAATA 2034
DB 223414 AAGATTAATCTGGAATTTCAAAATTAATCTTTGCTAATTTAAATGTAAGAAATTAATC 223473
QY 2035 GAGGATATGTTCAAGAGAGAGACCCAGAGTCAAGTGTGTTTGAAGGTTGAAGAGAAAC- 2093
DB 223474 TCAAAATCTCTAAGATAGTTTACATGATGAGCAATTAATGATTAAGTCACTGCAAAA 223533
QY 2094 CAGTGTGTTGCAATCTCCAGGTTTCAATGATGCTGTTTGAATTAATGATGATGATGATG 2153
DB 223534 CATTAATTTCTTTAAATATTTCTGCTTAATGATTAATTAATTAATTAATTAATTAATTA 223593
QY 2154 ACCGTGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2213
DB 223594 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 223649
QY 2214 TTAATTTGAAGAGCAAGATTAATTTTCTCAACTTTGCTCAATTTGAATCTGATCT 2273
DB 223650 TCACTTTGAAGAGTCAATTTGAATTAATTTAG-----TGTTAACAAAGCA 223695
QY 2274 TATATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2233


```
Db 222968 CAATATCAAGAGTGCAGAGAAAGCT-----CACTTAAAGATTAATTAATATATTA 223023
QY 1560 GAAATACGATCTAGTACTAGATGTGCACACGACATCAAGATATCTTCTAGTTTCAAGA 1619
Db 223024 CAATGTGTCACTATGTATTTATCTGAGGGAGTTAATACAGAGAAATGTTCAAAATTA 223083
QY 1620 AGTTTCAGATCGGCGCGGCGGTGTCTACGCTGTAAATCCAGACCTTTGGAGGCGG 1679
Db 223084 AGTAACATA-CAGCCGGGCAATGGGTCTCATGCTTAATCCACACCTTTGGAGGCGG 223142
QY 1680 AGCGGGGTGATCAC--AAGTCAGAGATCAAGACCATCTGATTACACGGTGAACC 1737
Db 223143 AGCGAGCGGATTAATCTGAGGTACAGACTTTGAGACCAAGCTTAACCAACTTTGTAACC 223202
QY 1738 CCGTCTCTACAAAATATATATAAAATTTAGCCAGCGGTGTGGCGGCGCTGTAGTCCA 1797
Db 223203 CTGTCTCTACTAAATATGTATATAAAATTTAGCCAGCGGTGTAGTGGCGGCTGTAAATTTCA 223262
QY 1798 GCTAATCAGAGGCTGAGGAGAGAAATGGCGTGAACCCCGGAGTGAAGATTGCG-GTGA 1856
Db 223263 GCTAATCAGAGGCTGAGGAGAGAAATCTTGAACCCAGAGGAGGAGAGGTTGCACTGA 223322
QY 1857 GCCGAGATCGCGCACTGCGCTCCAGGCTGG--CGACAGATGAGACTGCTCAAAA 1914
Db 223323 GCTGAGATCATGCTATGCACTCCAGGCTGGGACACAGACAAAATCTCCCTCTCAAAA 223382
QY 1915 AAAAAAAAAAAAAAAAAAAGTTTCAAGTCTTAAACACTGATTTCAACAGCTTAGAA 1974
Db 223383 AAAAAAAAAAAGTAACA-----TCATACAGCGTAAC 223413
QY 1975 TAGAGAGCATGTATACAGGAGAGAAATGTTTTCAGCAAGATACAGATAGGAAAT 2034
Db 223414 AATATTAATCTGATGATGTCAAAATATATCTTTGCTACTTTTAAATTTGAAATTAATC 223473
QY 2035 GAGGATATGTTTCAAGAGAGAGACCCACAGATCATGGTTTGTAGGGTTAGAGAAAACA- 2093
Db 223474 TCAAAAATCTCGAAGATAGTTACTGATGSCATTAATATGATTAAGACCTGCAAAAATAG 223533
QY 2094 CAGTGTGTTGCAATCTCCAGGTTCCATTAAGCGGTTATGAATCAATATGTTGGTTAGCA 2153
Db 223534 CATATTTTCTTTTAAAAATATCTGCAATTAATATTAATTAATTAATTAATTAACAAACAC 223593
QY 2154 ACCTGATTTTAAAAATGAATTAATTAATGATGAGAGAGAAATGAATTTTGCATGCA 2213
Db 223594 ACATACATATGTATATGTGTACTATGTTTATGATATACCTTCTCCATTT---TATA 223649
QY 2214 TTACATTTGAAAGAGCAAGTATATTTTCTGCACTTTTGTCAATGTATGTACTGACT 2273
Db 223650 TCACTTGTAAAAGTCCACATGTGAAAATAG-----TGTACACAAAGCA 223695
QY 2274 TATATTTTATGATGATGTGATGATACAGATACATATATTTTCTTACTGTAGCTGCA 2333
Db 223696 AAGAAATACAGGCTAGCTAGGAGGAGCTAGCTATTAATTTTCTTATTAAGCATGATCTTG 223755
QY 2334 GTCAAAAATCTTTAAGCACTGGCTGTGCTAACTTCTTATTTTTCAGAGAGAAATC 2393
Db 223756 GACAAATACCATCTCTGTGACCAAGCTTCTGTTTCTTATCTGTGAGATGGGGCA--- 223812
QY 2394 CAAGATCTGAGAGACAAACATTTTCCGTAGGTTATAGAACCAAGCTTATGCAATTTGCTA 2453
Db 223813 -----TATATGACAAACCTTGGCATCCCATTTTACCTTAAGGTACATPAAGAT 223864
QY 2454 AAGGTATTTTGTATTAATTTCTTTCCACTAGTGTGCATACGCACTTCTAGTTCTGTT 2513
Db 223865 ACCCTAAGGTAAATATAAATATATTTGCAAGGTATTAATTTCAATATGATCTTAT 223924
QY 2514 GGCTGAAATATACAGATATATATGTAAGACAGATACACAGTCTGGGAGAAATATATG 2573
Db 223925 CAACACGATGTCTTAATTTT-----TATTTCTATGATTAATAAATG 223969
QY 2574 GTAGGTGCTGAGGCTCAATTTTCTAAGAAATGTGAACTTATGAGGAGGTATGTGGCT 2633
```

```
Db 223970 CAGACCAATTTGAAATCTCAGGTTGACGTCAAGATTAATCTCCCGGCCAAGCACATGGCT 224029
QY 2634 CACACCTATTAATTCAGACCTTTGGAGGCCAAGTCAAGATCGCTTGAACCCAGAG 2693
Db 224030 CACAACTGTATTCGCCAGCATTTTGGAGGCTGAGGAGGCGACTCATTTGAGGTCAAGAG 224089
QY 2694 TTCAAGATCAAGATGGGCAACATAGACACCTCATCTCTTCAAAAATTTTAAAAATCG 2753
Db 224090 TTGGAACCAAGCTTGACAAATGTGAAACCCCATCTCTTCAAAAATTAACAAAATTTAG 224149
QY 2754 CTGAGCATGTGGCATATACGCTGTAGTCCCACTTACTCTGGAAGCTA--GTTGGGTGATC 2812
Db 224150 CCGGCGATGTGTGGACGTGTGCTGTAGTCCAACTGTTTGGAGGCCATGGAGGAAATC 224209
QY 2813 GCTTG-ACACAGAGTTTGAAGCTTAAGGCTAAGGCTATGATCAACAACTGCACTCAAGCTT 2871
Db 224210 ACTTGAACCAAGAGATGAGGTTGCAAGTCCAGATCGACATGCACTCGAGCT 224269
QY 2872 GAGTGCAGAGAGAACCTGTCCCTTAAAAAAGAAAGAA 2911
Db 224270 GGCGAACAGAGCAAGCTCCTCTCAACAAAAA 224309
```

RESULT 11

```
US-09-814-353-20332
; Sequence 20332, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US-09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 20332
; TYPE: DNA
; LENGTH: 2824
; ORGANISM: Homo sapiens
US-09-814-353-20332
```

Query Match 5.5%; Score 338; DB 10; Length 2824;

Best Local Similarity 88.0%; Pred. No. 1.3e-67;

Matches 368; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

```
QY 16 CAATTTAAGAAAGCAAAATATACAGCTTAAATAGAAATGTGAGACTAAATACCAAG 75
Db 332 CAATTTAAGAAAGCAAAATATACAGCTTAAATAGAAATGTGAGACTAAATACCAAG 391
QY 76 AACTCCCTTTGATGATTTGTAAACAAAATTAATATGATGATGATTAATGTTCTAATG 135
Db 392 CACCCCTTTGTATGATTTGTAAACAAAATTAATATGATGATGATTAATGTTCTAATG 451
QY 136 GAATGTGAACCAAGACCATATCAGGCTTAGCAAAATGGCAATTCATATATATCA 195
Db 452 GAGTGTGACCAAGACCATATCAGGCTTAGCAAAATGGCAATTCATATATATCA 511
```

QY 196 AGATTCTCCTTCAAGAGCTTCAGGCGCTAATGATCTCTAAGAAATGTAAGCCCTC 255
DB 512 AAGTTCTCTGCAAGAGCTTCGCGCTAATGATGCTAAGAAATGTAAGCCCTC 571
QY 256 AGGCATCTGAAGAGAGCTGTTACAGCAATTCATCAAAAAAGAAACACAGCCCTTCC 315
DB 572 AGCGCGCGAAGAGAGCTGTTACAGCAATTCATCAAAAAAGAAACACAGCCCTTCC 631
QY 316 CTTCCTCCCATCTGATGTAAGAGCTTCCTTCATTCATGTAATTTCTTAATAC 375
DB 632 CTTCCTCCCATCTGATGTAAGAGCTTCCTTCATTCATGTAATTTCTTAATAC 691
QY 376 AGCTGTAGAGCTCAAGTAAGTAAGAGCTCCCATTCAGAGAAATTTATCTTA 433
DB 692 GTCTGTAGAGCTCAAGTAAGTAAGAGCTCCCATTCAGAGAAATTTATCTTA 749

RESULT 12
US-10-002-491-10
Sequence 10, Application US/10002491
Publication No. US20030109467A1
GENERAL INFORMATION:
APPLICANT: Brett P. Monia
APPLICANT: Andrew T. Wale
TITLE OF INVENTION: ANTISENSE MODULATION OF FXR EXPRESSION
FILE REFERENCE: RTS-0239
CURRENT APPLICATION NUMBER: US/10/002,491
CURRENT FILING DATE: 2001-11-15
NUMBER OF SEQ ID NOS: 88
SEQ ID NO 10
LENGTH: 91000
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 73772-74071
OTHER INFORMATION: n = A, T, C or G
NAME/KEY: Intron
LOCATION: (576) ... (19018)
OTHER INFORMATION: Intron 1
NAME/KEY: Intron
LOCATION: (19154) ... (19785)
OTHER INFORMATION: Intron 2
NAME/KEY: Intron
LOCATION: (19919) ... (37293)
OTHER INFORMATION: Intron 3
NAME/KEY: Intron
LOCATION: (37660) ... (58973)
OTHER INFORMATION: Intron 4
NAME/KEY: Intron
LOCATION: (59115) ... (61405)
OTHER INFORMATION: Intron 5
NAME/KEY: Intron
LOCATION: (61540) ... (63027)
OTHER INFORMATION: Intron 6
NAME/KEY: Intron
LOCATION: (63127) ... (63463)
OTHER INFORMATION: Intron 7
NAME/KEY: Intron
LOCATION: (63564) ... (67187)
OTHER INFORMATION: Intron 8
NAME/KEY: Intron
LOCATION: (67335) ... (87922)
OTHER INFORMATION: Intron 9
NAME/KEY: Intron
LOCATION: (88037) ... (89288)
OTHER INFORMATION: Intron 10
US-10-002-491-10

Query Match 5.5%; Score 338; DB 15; Length 91000;
Best Local Similarity 45.5%; Pred. No. 8, 2e-67;
Matches 1134; Conservative 0; Mismatches 1280; Indels 78; Gaps 13;

QY 471 ATATATAGTTGTTAGCCGTGGCGGTGAGCTGATCCCTGTAATCCAGGCCCTTGGGAG 530
DB 72916 AATTCACCTTCATTCGCGGGCGGTGAGCTGATCCCTGTAATCCAGGCCCTTGGGAG 72975
QY 531 CCAGAGTGGGAGATCATGAGTCAAGATTGAGACCAAGCCATGACCAATGTTAAAC 590
DB 72976 CTGAGCAGTGGATCAAGAGTCAAGATTGAGACCAAGCCATGACCAATGTTAAAC 73035
QY 591 CCGGTCTCTAATGAATCAAAAAATTAGCAGTGTGTGGCGCAGCACTGTAAACCC 650
DB 73036 CTGTCTCTAATGAATCAAAAAATTAGCAGTGTGTGGCGCAGCACTGTAAACCC 73095
QY 651 AGCTGTCAAGAGAGTGAAGAGAGAAATTCCTGTAACCCAGAGGCAAGG--TGAAGG 709
DB 73096 AGCTACTGAGAGGCTGAGGAGAGAAATTCCTGTAACCCAGAGGAGTGAAGGCAAGT 73155
QY 710 AGCAAGATCACACCAATGACTGTAGCTGTGATGACAGGCAAGACTCAATCAAAA 769
DB 73156 AGCCAGATCTGCGCACTGCACTCCAGCCTGGGTGACAGAGAGCTGTCTCAAAA 73215
QY 770 AAAAAAAAAAAGAAATGTAGTTGTCTAT-----AACAAATTA 812
DB 73216 AACAAACAAACAAACAAACAAATCAAAATCATTTGTGACCTTCATCCACTTA 73275
QY 813 ATAGGAGTGA--GAAGCAAGTCTAAGCCTATGACCATGTACTAGAACTACTGTG 870
DB 73276 CCTTCAGCATTGACTTACTCTCTTGTGCTGACCACTTCTCTAGTCTTGAGCTT 73335
QY 871 GGAACATATATAGGAACTTAACCACTGCTGAA-----GTAAAGT 914
DB 73336 AGAGTCTGTGTTCACTCAATATGATATCTTCTTGTAGATCATTCACAGATATCTCA 73395
QY 915 TTTGAAAGAAATGTTTGAAGCAAAAGGTTAAAGAGTGAAGAAATTAATTAATAC 974
DB 73396 CTTTGCTTTCTTCACTTAAGAAACCTTGCGCTGAGTGAAGTCTTAACTTAA 73455
QY 975 CAGTTAGCTGTGTGAGAAATGGATAGGAGCTAATGAGAAATCAATAGGAATGTT 1034
DB 73456 TCTGTGCTTGAAGCTGAGAACTGAACAGATGAGTGAAGTCAACCAAGTTAGTATCA 73515
QY 1035 TCATGTATGTTAAGAACCTGGTAAAGGTGAAGACATTAATCTGACCATTCGCG 1094
DB 73516 AACTTCAAGTGAAGCTCACTGATCTTCCATGATCTCTTTCTTGAAGCAATTCATCTT 73575
QY 1095 GGAATTTTTTATGTAATGCTTGGCAATTT-----AATAGAGAGCA 1140
DB 73576 ATACTTATGAAATATGTTGATGCTTCTTCTTCAAAATTCACATCTCTTCCA 73635
QY 1141 GAGATGTGACAGTTGATGATGATGAGTTGAAGTTCTGCCAGACATGTGAAGAA 1200
DB 73636 CTCAATCTCACTATTAATCTTGTGAGAAAGTGTCTTCAAGACATGCAATGAGAT 73695
QY 1201 AGACAGTGAAGAGAGTGA-----GAGATTCAAGACAGAAATATGTGTGCG 1256
DB 73696 TGAAGCTCAAGATATGTGTCAACCCAGCTTCAAGAAATGTAATCAAGCTTGAAG 73755
QY 1257 CAGTGCATCTAGTGTGATCTAATCTGAGGAGAGAAATGAAGTAAACAGCTTGTGA 1316
DB 73756 AAGATGAATCCCGAGNN 73815
QY 1317 TTAGTTATGAAGAGTGAAGGCTTCAGAGCTTACAGGTGTGATTAATGAAGATG 1376
DB 73816 NNN 73875
QY 1377 ATTGAGAAAGATTAATCTGAGAGAGTATTTCAAGCTTGAAGTCACTCATAC 1436
DB 73876 NNN 73935
QY 1437 CAGACACTGTGCTTAATGCTTCAAGACATGATCCCTGCTCAAGGAGCTTACAGCCA 1496
DB 73936 NNN 73995
QY 1497 AAACAAGAGATTAAGAAATACACCAATATCTTATAGACACTGTGTGATGATCAAG 1556

Qy	777	AAAAAGAAATNTAAGTGTGCTATTAACAATAATTAAGCACTGAGAAACAAACTGCT	836
Db	13420	AAAAAAAAAATCTGTAATGGAAGACAT--CAGTATATTAGTACTTTAAAGACATATATT	13363
Qy	837	AAAGCCTATGACATGTGTAATAGAAATTAAGTGTGGAAACAATATTAAGGAACCTTACC	896
Db	13362	AATGAGAAAAACAGCTATTAAGAATTAATAGCATTTGTATCTTACATTAAGAA-----C	13309
Qy	897	CAGTCTTGGAAGTAAGTTTTGGAAAGGAATGTTTGAAGCAAAAGGTTTAAAGAGTGA	956
Db	13308	AAGAAAGTAAATTACAGGTGAAGATCTAAATATTTTAAAAATTTAAATCTGAAGAAAT	13248
Qy	957	AAAAAAAAAATTTAAATACAGTTTAGCTGTGTGGAGATGGAGTAGGAGCTAACAGAG	1016
Db	13248	ATAGAAATATATTTTGGACAGATCTTTGGACAGAGAGATCTTCCAAAGCATGATATGAAT	13188
Qy	1017	AAATCAAAATAGGAATGTTTCAATGTAATGTTAAGACCCTGTGAAGGTGAAGCAATTAC	1076
Db	13188	TCAAAACCCAAATTTTACTTCAAGAAATTTTCTCAATGGGAAAAAGACAACTTAA	13128
Qy	1077	ATTATCTGCACCATCGGGGACTTTTTTTTATGTATATGCTTTGGCAATTTTAAATGAGG	1136
Db	13128	ATTAAATGCTACATTAGTAACAACAAAGTTAAATTAATCTTAAACACATAGAGCTACT	13068
Qy	1137	AGCAGAAATGTAGACAGTTGATGATGAGTCAGATGGAAGTTGCAATCTGCAGACATGTAAAG	1196
Db	13068	CCAATTACTAAGAAAAATTTTAAAACTAAAGAAAA--TGCACAGATTAAGACTAG	13012
Qy	1197	GAAGAGCAGGTAGGCAAGAGATTTGAAGATTTATCAGACAGAAATTAATGTGCTGCG	1256
Db	13011	CAAGTCCAG-----AAGAAATAATATGACTTAACAAATAGAGAGCCGTGTGAGTA	12960
Qy	1257	CAGTGCATCTAGTCTGATGCTAATCTAGAGGAAAGAAAGTAAGATTAAGACGCTTCTGA	1316
Db	12959	TTACGCGCTGTAAATCCAGACATTTGGAGGCGGAGCAGCAGATCACTGAAGTTGGG	12900
Qy	1317	TAGTTATGAAGAGTGAAGGCTTCAAGACCTACAGTGTGATTAATTAAGAGATG	1376
Db	12899	AGTTCAAGACAGGCTGATCCAAATGAAAAACCCACTCTCTCTAAATAATCAAAATTA	12840
Qy	1377	ATTG-----GAGAAAGAAATTAACGTGTGAAGATGAGATTTTCAGGCTTGAGTGA	1431
Db	12839	GCTGAGGTGTGTGGCGCATGCTGTAAATCCAGCTACTCCAGAGGCTGAGGCAAGAAAT	12780
Qy	1432	CATACAGACACTGTGCTAAATGCTTCAAGACATGATCCCTG-----CCCTCA	1480
Db	12779	CGCTTGAACCCGTGAGGCAAGGTGCAATGAGCAAGATCTCGCCACTTGCACTCCAGC	12720
Qy	1481	AGGGACTTAAGCCAAATAAAGATTAAGAAATTAACACCAATACTATTTTGAACACT	1540
Db	12719	TGGGCAACAGAGTGAATCTGTCTCAAAAAAACAACAAAAACAAAAAGAAAG	12660
Qy	1541	TGTTAGAAATATCAAGAAAGAAATACGATTAAGTACTGTGATGTCAACGCGATCAAG	1600
Db	12659	TGCCACACCTCAACCAAGAAATAGAGAAAAATGAATTTAACAATGTACAGTACTTTT	12600
Qy	1601	ATATCTTCTAAGTTCAAGAAATTTCAAGATGCGCCGGCGCGGTGCTCACGCTGTATC	1660
Db	12559	ATCAAAAAAATAGT-AATTAATAATTAAGAGGCGCAGGCGCATGTGATCTACGCTGTATTC	12541
Qy	1661	CCAGCACTTTGGGAGGCGGAGGCGGGGTGATCAAGAGTCAAGAGATCAAGACATTCCTG	1720
Db	12540	CCAGCACTTTGGGAGGCGGAGGCGGGGTATCACAGGTCAAGGGAATCAAGATCATCTTG	12481
Qy	1721	GTTAACGGGTGAACCCCGTCTCTACAAAAAATATTAATAATTAAGCAGGCGTGTGCG	1780
Db	12480	GCTAACATGTGTAACCCCGTCTCTACTAAATAATCAAAAAATTAACCGGCGTGTGTGT	12421
Qy	1781	GGGGCGCTGTAGTCCCAAGCTACTCAGAGGCTGAGCAGAGAAATGGCGTGAACCCGGGA	1840
Db	12420	GGGGCGCTGTAGTCCCAAGCTACTCAGAGGCTGAGCAGAGGAATGCGGTGAACCCGGGG	12361

```

QY 1841 GGTGAGATTGTC-GTGAGCCGAGATCGGGCCACTGCGCTCCAGCTGGGGCAGACAGATGA 1893
      |||||
Db 12300 AACGAGATTGCACTGAGCCGAGATGTGGCCACTGCACTCCAGCTGGGAGACAGAGCA 12310
QY 1900 GACTGCGCTCTCAAAAAAAAAAAAAAAAAAAAAA 1933
      |||||
Db 12300 GACTCCGCTCAAAAAAAAAAAAAAAAAATATATATTA 12267

RESULT 14
US-09-764-891-8149/c
; Sequence 8149, Application US/09764891
; Publication No. US20030077808A1
GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8149
; LENGTH: 31474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-8149

Query Match 5.4%; Score 334.4; DB 10; Length 31474;
Best Local Similarity 57.2%; Pred. No. 3,2e-66;
Matches 755; Conservative 0; Mismatches 536; Indels 29; Gaps 7;

QY 1607 TCTAGTTTCAAGAGTTTCAGATCGCCGCGCGGTGGCTCAGCCCTGTAATCCAGCA 1666
Db 1313 TGTCACTGTAAATAATATATTTTTCAGCCGGGCGCGGTGCTCAGCCCTGTAATCCAGCA 12544
QY 1667 CTTTGGGAGCGCCGAGCGCGGTGATTCACAAAGTACAGAGATCAAGACATCTCTGTTAAC 1726
Db 1353 CTTTGGGAGCGCGGCGCGGTGATTCACAGAGTACAGAGATCAAGACATCTCTGTTAAC 11944
QY 1727 ACGGTGAACCCCGCTCTCTACAAATAATATATTAAGCCAGGCGTGTGCGGCGCGC 1786
Db 1193 ACGGTGAACCCCGCTCTCTACTATAATAATTAAGCCGGCGTGTGTAGCGGCGCGC 11344
QY 1787 CTGTAGTCCCACTACTCAGAGGCTGAGGCAAGAGATGGCGTGAACCCGCGAGGTAGA 1846
Db 1133 CTGTAGTCCCACTACTCGGAGGCTGAGGCAAGAGATGGCGTGAACCCGCGAGGCGGA 10744
QY 1847 GTTTTC-GTGAACCGGAGATGCGCCACTGCGCTCAGCCTGGGGCAGACAGATGACTGC 1905
Db 1073 GCTTGCAGTGAAGCCGAGATCGCGCACTGCACCTCGGGGAGACAGACGAGACTCC 10144
QY 1906 GTCTCAAAAAAAAAAAAAAAAAAAAAAGTTTCAGATCTTAACACACTGCATTTCAAC 1965
Db 1013 GTCTCAAAAAAAAAAAAAAAAAAAAAAGTTTATATATATATATATATATA 954
QY 1966 AGTCT-----AGATAGAGAGAGCATGTTACAGGAGAGAAAAAGTTTTCAGCAAGAGTA 2019
Db 953 TATTTTAAACATATATAGAGAGATACAAATGTTTCAATATCTGTTTGTCTAATAACT 894
QY 2020 CAGAGTGGGAAATAGAGATATGTTTCAGAGAAAGAGACCCACAGATCATGTTGTTAG 2079
Db 893 TGGTTGATGATACCAAAATATGTAATTTTGTGCTGTTTGTCTAATTTGTCTAG 834
QY 2080 GGTTAGAGAAACACAGATGTTTTCATCTCCAGGTC-----CATTAGTGCGTATGAA 2135
Db 833 TCATTAAGCATTAAGAACTTTGATCTTAACGTGCAATGTTCTGGAACCCAGACAGATACATAG 774
QY 2136 TCAATATATGTTAGTACCACTGATTTTAAAAAATGAATTAATGATGATGAGAGAGAAAT 2195
Db 773 AGAAGCAATATATCTTAAGGCTTTTAACCTTAAGAGCTTGGGAAACCTGGAATTTCTGTGTC 714
QY 2196 AGAAATAATTAAGCATTCATTTATTTTGAAGAGCAATATATTTTCTGCACTTTTCG 2255

```

Db 713 AGATTTAAATATGAGAAAACATAAAGGTCAGGTGGCTTATGTTCCATGACTAC 654
QY 2256 TCCAAATGTAACGTACTTATATTTATGATGATGATAATCCAGATACATATATAT 2315
Db 653 CCACACAGTAAACAAGTTCACTTACAAATTAACGAAATAATTTAATCCAAAAGCTGCC 594
QY 2316 TTCTTACTGTAAGACTGCACTCAAAAATCTTTAAAGCATGCGCTGCTTAATCTCTTA 2375
Db 593 ATGATCCCAAAATTTTTCAGAGAAAATGACAAGGGAGTTGGCTACAGTTAGGACTAA 534
QY 2376 TTTTGCAGAGAGAAATCCAGATCTGAGAGACAAATTTTGCAGAGTTATGAAAC 2435
Db 533 GAATGAACATGTAAGACTGAATCTTAACCCCTTAAATTTACATGAATTAAGAAAA 474
QY 2436 CAGCTTATGCCATTCCTAAAGATGATCTTATTAATTTCTTCCACTAGTCCATAC 2495
Db 473 AATTTCCTCAAAATGTAACATCAATGAAAAAATTAATTCCTCAATTTCAAACTAC 414
QY 2496 TGCACCTTCTAGTTCTGTGGCTGAAATCAGAATATATTAGTAAACAGCATACAG 2555
Db 413 AGAGTTCCTTTTATTTATTTGCAACACTTCACTTACATATAAATTAATAAAAACTG 354
QY 2556 TCTGGGAAATATATTTGGTAGAGGTGCTGAGAGCTGATTTCTTAAGAAATGGAAGCTT 2615
Db 353 CATGAAACATTAAGATGCGCTATAG-----TGTATCATTAATAAATTTAATTT 306
QY 2616 AGCAGAGATATGTTGCTGCTACACCTATATATTCAGCATTTGGAGGCCAAGTCAAGAA 2675
Db 305 AGGCTGGGCGCGGTGCTCAAGCTGTATCCAGGACCTTTGGAGGCCAGGCGGGCGG 246
QY 2676 ATGCTTGAACCCAGAGTTCAGAGCTAGATGAGGCAACATAGCAGAACTCTCATCTTAC 2735
Db 245 AT--CATGAGTCAAGAGATGAGAACCATCTGCTTACACGATGAACCCCTCTCTAC 188
QY 2736 AAAAAATTT---AAAAATCAGCTGAGCATGTGSCATACGCTGTAGTCCCACTACTG 2792
Db 187 TAAAAAATACAAAAAATTAAGCCGGGCGTCTGCGCGGCACTGTAGTCCCACTACTG 128
QY 2793 GGAAGCTAGTGGTGTGATGCTTG--ACACAGAGATTGAGGCTTAAGTGAGCCATGATC 2851
Db 127 GAGGCTGAGGACAGAGATGCGTGAACCTGGAGAGCGGCTTGAGATGAGCCGAGATT 68
QY 2852 ACACAATGCACTCCAGCTTGAAGTGAACAGAGAAAGCCGTGCTCTTAATAAAGAAAGAA 2911
Db 67 GTGCCACTGCACTCCAGCTGGCAACAGCAAGACTGTCTCAAAAAAAGAAAAA 8

RESULT 15
US-10-723-860-1627/c
; Sequence 1627, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1627
; LENGTH: 111084
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-723-860-1627

Query Match 5.4%; Score 334; DB 18; Length 111084;
Best Local Similarity 55.1%; Pred. No. 7.8e-66;

Matches 803; Conservative 0; Mismatches 640; Indels 15; Gaps 7;
QY 483 TTAGCTGGCGGGTGGCTCAGCGCTGTATCCAGCCCTTTGGAGGCCAGAGTGGCA 542
Db 50920 TTGGCCAGCGGAGTGGCTCAGCGCTGTATCCAGCACTTTGGAGGCCAGAGTGGCG 50861
QY 543 GATCATGAGTACAGAGTTTGAACCAAGCTTACCAATGCTGAAACCCGCTCTACT 602
Db 50860 GATCAGAGAGTACAGAGATGAGACCATCTGCTTACACAGGTGAATCCGCTCTACT 50801
QY 603 AAAAGATCAAAAAATTAAGCAGGTGTGGGAGCAACCTGTAAACCCGCTGCTGAG 662
Db 50800 AAAAATACAAAAATTAAGTGGGCGGTGAGTGGGCGGTGCTGTAGTCCAGCTAC 50741
QY 663 GAGTGGAGGAGAGATTTGCTTGAACCCAGAGAGAG--GTGACAGCAAGATCAC 721
Db 50740 GCGTAGGAGAGAGATGCGTGAACCGGAGGAGGAGCTTACAGTGAAGCGAGTAGC 50681
QY 722 ACCAATGCACTGAGCTGATGACAGGAGCAAGACTCCAACTCAAAAAA 781
Db 50680 GCCACTGCACTGCGCTGGGGAAGAGCGAGCTCCATCTCAAAAAA 50621
QY 782 AGAAATATGTAAGTGTGCTATACAAATTAATAGGCAAGAGCAAAAGCTTAAGC 841
Db 50620 A--AAGAGGTGAATTCGCTTAGGCAAGGTGACTCACACTGTATTCAGCACTG 50562
QY 842 CTATGACATGTGTAATAGAAATCTGTGGAAACATTAATAGGAACTTAACCACTC 901
Db 50561 GGCAGAGGTGGCGGATGTTGAGCCAGAGGTTGAGAACTATCTGCTCAACAGAG 50502
QY 902 CTGGAAGTAAAGTTTGAAGAAAGATGTTGAGGCAAAAGGTTAAAGAGTGA 961
Db 50501 AGACCCCACTCCTCAATAAATAAATAAATAATTAAGTGGGTTGGCTGAGGCTC 50442
QY 962 AAAAAATTAATACAGTTAGCTGTGTGGAGAGATGAGATG--GAGCTAATAGAGAAAT 1020
Db 50441 AGCGCTGTATCCAGCACTTTGGAGTGTAGGGGGGTGATCATCTGAGGTCA 50382
QY 1021 CAATAGGAATTTTCAATGATGTTAAGAACCTGTTAGGGTGAAGCCATTACATTA 1080
Db 50381 TCAAAACCGCTGACCAACATGAAAAAACCCTCTACTAATAAATAAATAATGCTG 50322
QY 1081 TCTGACCATGCGGGGACCTTTTATTGTAATGCTTGGCAATTTAAATGAGAGAGA 1140
Db 50321 GCGGTGTGGGCGCATGCTGTATCTTACTAGCTAGGAGGTGAGGAGGAAATGTG 50282
QY 1141 GAGAAATGTAACAGTTGATAG-----TCAGAGTTGAAGTTCTGCCAGACATGTG 1192
Db 50261 TGAACCCAGAGAGACGAGGTTGCGATGTGCCAAGATGTGTCTACCTCCAGCTGGG 50202
QY 1193 AAGAGAGAGAGATGAGCAAGAGATTTGAAGATTAATCAAGACAGAAATTAATGTGC 1252
Db 50201 CAAGAGAGCAAACTGCGTCTCAAAAAAAGGCTGGGCAAGTGTG 50142
QY 1253 TGCGCAGTGGCA--TCTAGTCTGAGCTTATGTGAGGAGAGAGATTAAGACAGCT 1310
Db 50141 TGGGCGCTTTGATTTCCAGTACTAGAGGCTGAGACAGAAATTTGCTGAACAGAG 50082
QY 1311 TGCTGATAGTTATGAAGAGAGTGAAGCTTCAAGACCTACAGTGTGTTAATAG 1370
Db 50081 AGCAGAGAGTTGCAAGTGGCCAAAGATGTGCCATTACAGCTGGGTGACAGAG 50022
QY 1371 AGAATGATTTGAAGAAATTAATCTGTGAGAGTGAAGTTTTCAGGCTTGAGTACTTC 1430
Db 50021 CAAAATTCATCTCAAAAAAAGTAACTGGGATGAGTGTGCAATGTGTCTGTACTC 49962
QY 1431 ACATACAGACACTGTGCTTAATGCTTCAAGACATGATCCCTGCAAGAGCTTAC 1490
Db 49961 CCACTACTCAGAGAGCTGAGGTGGAGAAATTTGCTTGAAGAGAGAGTCAAGCTGTAG 49902
QY 1491 AGCCAAAAACAAGATTAAGAAATACACCAATATTAATTAAGACCTTGTGT--AGAA 1549
Db 49901 TGAGCCATTAATGTGCTCACTCAATCCAGCTGGGCAACAGATGAGACCCCATCTCA 49842

```
QY 1550 TATCAAGAAAGAAATGCGATCTAGTACTGTAGATGTGCAACGGCATCAAAAGATATCTTCT 1609
Db 49841 AAAAAAAAAAAAAAAAAAAGAGGTGAACCTGAGTTCACTTGCTTGTGAGAGACAAGAA 49782
QY 1610 AGTTTCAAGAAGTTTCAGATCGGCGGGGGCGGTGCTCACGCTGTATATCCAGACATT 1669
Db 49781 GTAGCAAAAGTTGGGTGTAAACAGCCAGGTGTGTGGCTCACACATGTATATCCAGACATT 49722
QY 1670 TGGAGGCGGAGGCGGGGTGATCAAAAGTCAAGAGATCAAGACATCTGTGTTAACAG 1729
Db 49721 TGGAGGCGTGAAGCTGGCGGATCAGAGGTCAAGAGATTGAACATCTGTGTTAACAG 49662
QY 1730 GTGAAACCCCGTCTTACAAAAATTAATAAAATTAGCCAGCGGTGTGGCGGCGCTG 1789
Db 49661 GTGAAACCCGTCTTACTTAAATAAATAAAGATTAGCCGCGGTGTGGCGGCGCTG 49602
QY 1790 TAGTCCAGCTACTCAGAGGCTGAGGCGAGAGATGGGCTGAACCGGGAGGTAGATT 1849
Db 49601 TAGTCCAGCTACTCAGTAGGCTGAGGCGAGAGATGGCGTGAACCGGGAGGCGAGCT 49542
QY 1850 TGC-GTAGCCGAGATCGCGCCACTGCGCTCCAGCTGGGCGACAGAGTGAGACTGCGTC 1908
Db 49541 TGCAGTAGGCGCGAGATGGGCGCACTGCACTCCAGCTGGAGCGACAGAGTGAGACTGCGTC 49482
QY 1909 TCAAAAAAAAAAAAAAA 1926
Db 49481 TCAAAAAAAAAAAAAAA 49464
```

Search completed: December 27, 2004, 12:15:28
Job time : 2025.28 secs

This Page Blank (uspfo)


```
FT /number= 3
FT replace(3673,G)
FT /*tag= k
FT variation
FT replace(3747,G)
FT /*tag= l
FT variation
FT replace(3788,G)
FT /*tag= m
FT exon
FT 7939. .8018
FT /*tag= n
FT /number= 4
FT intron
FT 8019. .18260
FT /*tag= o
FT /number= 4
FT variation
FT replace(8034,T)
FT /*tag= p
FT exon
FT 18261. .18355
FT /*tag= q
FT /number= 5
FT intron
FT 18356. .19116
FT /*tag= r
FT /number= 5
FT exon
FT 19117. .19179
FT /*tag= s
FT /number= 6
FT intron
FT 19180. .29296
FT /*tag= t
FT /number= 6
FT variation
FT replace(27740,C)
FT /*tag= u
FT variation
FT replace(27752,T)
FT /*tag= v
FT exon
FT 29297. .29420
FT /*tag= w
FT /number= 7
FT intron
FT 29421. .31091
FT /*tag= x
FT /number= 7
FT variation
FT replace(29927,T)
FT /*tag= y
FT variation
FT replace(30772,G)
FT /*tag= z
FT exon
FT 31092. .31212
FT /*tag= aa
FT /number= 8
FT intron
FT 31213. .42825
FT /*tag= ab
FT /number= 8
FT variation
FT replace(36310,G)
FT /*tag= ac
FT variation
FT replace(36327,T)
FT /*tag= ad
FT variation
FT replace(40618,C)
FT /*tag= ae
FT variation
FT replace(40928,T)
FT /*tag= af
FT exon
FT 42826. .42919
FT /*tag= ag
FT /number= 9
FT intron
FT 42920. .50322
FT /*tag= ah
FT /number= 9
FT variation
FT replace(44701,A)
FT /*tag= aj
FT variation
FT replace(46020,C)
FT /*tag= ak
FT variation
FT replace(46036,G)
FT /*tag= al
FT variation
FT replace(46095,A)
FT /*tag= am
FT variation
FT replace(47608,G)
FT /*tag= an
FT exon
FT 50323. .50409
FT /*tag= ao
```

```
FT /number= 10
FT variation
FT replace(52150,A)
FT /*tag= ap
FT PN
FT MO200261060-A2.
FT PD
FT 08-AUG-2002.
FT XX
FT 17-JAN-2002; 2002WO-US001106.
FT PF
FT 31-JAN-2001; 2001US-0265151P.
FT PR
FT 09-MAR-2001; 2001US-00801861.
FT XX
FT (PEKE ) PE CORP NY.
FT PA
FT Yan C, Ketchum K, Di Francesco V, Beasley EM;
FT PI
FT P-PSDB; AAO18613, AAO18614.
FT DR
FT WPI; 2002-608515/65.
FT XX
FT P-PSDB; AAO18613, AAO18614.
FT XX
FT New human kinase peptide and nucleic acid molecule, useful for treating
FT PT disorders associated with abnormal expression of kinase protein, e.g.
FT adenocarcinoma of uterus or lung, in drug screening assays and
FT PT pharmacogenomic analysis.
FT XX
FT Claim 4; Fig 3; 131pp; English.
FT PS
FT XX
FT The present invention provides the protein, cDNA and gene sequences of
FT CC two splice variants of a human pftaire family kinase. The sequences are
FT CC specifically expressed in the human testis, brain, uterus endometrium
FT CC adenocarcinomas, lung fibroblasts, kidney renal cell adenocarcinomas, and
FT CC can be used to treat related diseases. The present sequence is the gene
FT CC of the invention
FT SQ
FT Sequence 53332 BP; 15976 A; 10335 C; 10952 G; 16069 T; 0 U; 0 Other;
FT
FT Query Match 100.0%; Score 6201; DB 6; Length 53332;
FT Best Local Similarity 100.0%; Pred. No. 0;
FT Matches 6201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT
FT QY 1 CACCATGGGCTCAGGCAATTGTAAGCCAAATATACAGCCTTAATATGATGTGAC 60
FT Db 20000 CACCATGGGCTCAGGCAATTGTAAGCCAAATATACAGCCTTAATATGATGTGAC 20059
FT QY 61 CTAAATACCCAGAGAACTCCCTTTGTAAGTTGTAACAAAATTATATGATGAG 120
FT Db 20060 CTAAATACCCAGAGAACTCCCTTTGTAAGTTGTAACAAAATTATATGATGAG 20119
FT QY 121 TTAATAGTCTAATGGAATGTAAGCCAAAGCCATACAGGCTAGCAAAATGGCAGA 180
FT Db 20120 TTAATAGTCTAATGGAATGTAAGCCAAAGCCATACAGGCTAGCAAAATGGCAGA 20179
FT QY 181 ATTCAATATCATCAAAAGTTATCTTCAAGAGCTTCAGGCGCTAATGATGCTAAAGAAA 240
FT Db 20180 ATTCAATATCATCAAAAGTTATCTTCAAGAGCTTCAGGCGCTAATGATGCTAAAGAAA 20239
FT QY 241 ATGTGAACGCCCTCAGCCATCTGAAGGACAGTGTTAACGAATGTATCAAAAAGAAAA 300
FT Db 20240 ATGTGAACGCCCTCAGCCATCTGAAGGACAGTGTTAACGAATGTATCAAAAAGAAAA 20299
FT QY 301 CCACAGGCGCTCCCTTCCCATATCTGATGTAAGCGTCTCAATTTCCATAGTAGT 360
FT Db 20300 CCACAGGCGCTCCCTTCCCATATCTGATGTAAGCGTCTCAATTTCCATAGTAGT 20359
FT QY 361 AAATTTCTAGATATACAGTGTGAGAGCTCAAGTAAGTGAAGAAAGCTCCCATTCAAAG 420
FT Db 20360 AAATTTCTAGATATACAGTGTGAGAGCTCAAGTAAGTGAAGAAAGCTCCCATTCAAAG 20419
FT QY 421 GAAATTTATCTTAAGATCTGTAATGATGTAATTTTGTACATTTGGATGTAATATAGT 480
FT Db 20420 GAAATTTATCTTAAGATCTGTAATGATGTAATTTTGTACATTTGGATGTAATATAGT 20479
FT QY 481 TGTTAGCCTGGGCGGTGCTCAGCGCTGTATCCAGCGCTTTGGAGGCCAGATGGG 540
```

Db	20480	TTGTAGCTTGGCGCGGTGGCTCAGCAGCTGTAAATCCAGCCCTTTGGAGGCGCAAGTGGG	20539
Qy	541	CAGATCATGTAGSGTCAGAGATTGGAGACACAGCCTTAGCCAAATGATGAAAACCCCGTCTTA	600
Db	20540	CAGATCATGTAGSGTCAGAGATTGGAGACACAGCCTTAGCCAAATGATGAAAACCCCGTCTTA	20539
Qy	601	CTAAAGATPACAAAAAATTAGCCAGGTGTGTGGCGCACCTGTAAACCCGAGCTGTGCA	660
Db	20600	CTAAAGATPACAAAAAATTAGCCAGGTGTGTGGCGCACCTGTAAACCCGAGCTGTGCA	20659
Qy	661	GAGGTGAGGCGAGAGAAATTCGTTGAAACCGAGAGCGAGGTCAGCGACGAAGATCA	720
Db	20660	GAGGTGAGGCGAGAGAAATTCGTTGAAACCGAGAGCGAGGTCAGCGACGAAGATCA	20719
Qy	721	CACCAATGSCACTGTAGCCTGGAATGACAGGGCAAGACCTCCAATCAAAAAAAAAAAAAA	780
Db	20720	CACCAATGSCACTGTAGCCTGGAATGACAGGGCAAGACCTCCAATCAAAAAAAAAAAAAA	20779
Qy	781	AAGAAATATGTAGATTGTGCTATPACAAATPAAATAGGCACTGAGAAACAAAGTCTPAA	840
Db	20780	AAGAAATATGTAGATTGTGCTATPACAAATPAAATAGGCACTGAGAAACAAAGTCTPAA	20839
Qy	841	CCTATGACCAATGTGTAATCTAGGAATACGTGTGGAAACACATPAAAGGAACCTPAA	900
Db	20840	CCTATGACCAATGTGTAATCTAGGAATACGTGTGGAAACACATPAAAGGAACCTPAA	20899
Qy	901	CCTGAGGTAGAGTTTGTGGAAAAGGAATGTTTGAAGCAAAAGGTTTAAAGAGTGA	960
Db	20900	CCTGAGGTAGAGTTTGTGGAAAAGGAATGTTTGAAGCAAAAGGTTTAAAGAGTGA	20955
Qy	961	AAAAATTPAAAAATACAGATTAGCTGTGTGAGAAATGGGATAGGAGCTPAACTAGAGAAAT	1020
Db	20960	AAAAATTPAAAAATACAGATTAGCTGTGTGAGAAATGGGATAGGAGCTPAACTAGAGAAAT	21019
Qy	1021	CAATATAGGAATGTTTCACTGTATGTGTAAGGACCTGTGTAAGGTTGAAGACATTAATTA	1080
Db	21020	CAATATAGGAATGTTTCACTGTATGTGTAAGGACCTGTGTAAGGTTGAAGACATTAATTA	21079
Qy	1081	TCTGCACCATCGCGGGGACTTTTTTTTTTATGTAATGCTTGGAATTTAAATACAGAGCA	1140
Db	21080	TCTGCACCATCGCGGGGACTTTTTTTTTTATGTAATGCTTGGAATTTAAATACAGAGCA	21139
Qy	1141	GAGAAATGTAGACAGTTGGATTGAGTCAGAGTTGAAGTTCTGCCAGACATGTGAAGGAAG	1200
Db	21140	GAGAAATGTAGACAGTTGGATTGAGTCAGAGTTGAAGTTCTGCCAGACATGTGAAGGAAG	21199
Qy	1201	AGACAGGTAGGCAAGAGAGTTGAAAGATTATCAAGCAGAAATTAAATGTGCTGGCCAGT	1260
Db	21200	AGACAGGTAGGCAAGAGAGTTGAAAGATTATCAAGCAGAAATTAAATGTGCTGGCCAGT	21259
Qy	1261	GGCAATCTAAGTCGTGCTAATCTGAGGGAAAGAAATGAATATPAGCAAGCTGTGTAATGT	1320
Db	21260	GGCAATCTAAGTCGTGCTAATCTGAGGGAAAGAAATGAATATPAGCAAGCTGTGTAATGT	21319
Qy	1321	TATATAAGAGATGGAAGGCTTCAAGGACCTACAGAGTGTGATTAAATAGAAATATGTTG	1380
Db	21320	TATATAAGAGATGGAAGGCTTCAAGGACCTACAGAGTGTGATTAAATAGAAATATGTTG	21379
Qy	1381	GAGAAAGATPACTGTGAGAGAGTGAATTTTCAGGCTTGAGTGAATCTCACTATCCAGA	1440
Db	21380	GAGAAAGATPACTGTGAGAGAGTGAATTTTCAGGCTTGAGTGAATCTCACTATCCAGA	21439
Qy	1441	CACGTGTCTAAATGCTTCAAAAGACATGATCCCGCTTCAAGGGAAGCTTACAGCAAAAC	1500
Db	21440	CACGTGTCTAAATGCTTCAAAAGACATGATCCCGCTTCAAGGGAAGCTTACAGCAAAAC	21499
Qy	1501	AAGAGATPAGAAATACACCAATACTATTATAGGACACTTGTGTAGAAATATCAAGAAAG	1560
Db	21500	AAGAGATPAGAAATACACCAATACTATTATAGGACACTTGTGTAGAAATATCAAGAAAG	21559
Qy	1561	AAATPAGATCTAGTACTGTAGATGTGCAACGGCATCAAGATATCTTTAGTTCAAGAA	1620

Db	21560	AAATACGATCTACTAGTACTGTACATGTCGCAACGGCATCAAGAAATATCTTCAAGTTTCAAGAA	21619
QY	1621	GTTCACAGATGAGCCCGGCGCGGTGGCTCAACGCTGTATATCCAGCACTTTGGAGGCCGA	1680
Db	21620	GTTCACAGATCGGCGCGGCGCGGTGGCTCAACGCTGTATATCCAGCACTTTGGAGGCCGA	21679
QY	1681	GGCGGCGTGCATCAACAGTCCAGAGATTCAAAGACATCCGTGGTTAAACAAGGTAAACCCCG	1740
Db	21680	GGCGGCGTGCATCAACAGTCCAGAGATTCAAAGACATCCGTGGTTAAACAAGGTAAACCCCG	21739
QY	1741	TCTCTCAAAAAAATATATPAAAAATTCACCAAGCGTGGTGGCGCGCTGTAGTCCAGCT	1800
Db	21740	TCTCTCAAAAAAATATPAAAAATTCACCAAGCGTGGTGGCGCGCTGTAGTCCAGCT	21799
QY	1801	ACTCAGAGGCTGAGGCAAGAGATGCGCTGAAACCCGGAGGTAGAGTTTGCCTGAGCCG	1860
Db	21800	ACTCAGAGGCTGAGGCAAGAGATGCGCTGAAACCCGGAGGTAGAGTTTGCCTGAGCCG	21859
QY	1861	AGATCGGCGCACTGCGCTCCAGCCTGGGCGCAAGATGAGACTGCGCTTCAAAAAA	1920
Db	21860	AGATCGGCGCACTGCGCTCCAGCCTGGGCGCAAGATGAGACTGCGCTTCAAAAAA	21919
QY	1921	AAAAAAAAAAAAAAAAAGTTTCAGATCTTAAACAACATCTTCAACAGTCTAGATTTAGAG	1980
Db	21920	AAAAAAAAAAAAAAAAAGTTTCAGATCTTAAACAACATCTTCAACAGTCTAGATTTAGAG	21979
QY	1981	AGCATGTTACAGGGAGAGAAATGTTTTCCAGCAAGATACAGAGTAGGAAAAATAGAGAT	2040
Db	21980	AGCATGTTACAGGGAGAGAAATGTTTTCCAGCAAGATACAGAGTAGGAAAAATAGAGAT	22039
QY	2041	ATGTTCAAGAGAGAGACCCCAAGTCAATGTTTTTTAGGGTTAGAGAAACAAGTT	2100
Db	22040	ATGTTCAAGAGAGAGACCCCAAGTCAATGTTTTTTAGGGTTAGAGAAACAAGTT	22099
QY	2101	TTGCAATCTCAGGTTCCATTAGTGGGTTATGAAATCAATATGAGTGGTTAGCAACTGCA	2160
Db	22100	TTGCAATCTCAGGTTCCATTAGTGGGTTATGAAATCAATATGAGTGGTTAGCAACTGCA	22155
QY	2161	TTTTTAAAAATGAATTAATGATGAGAGAGATGAATAATTAAGATGCACTTACAT	2220
Db	22160	TTTTTAAAAATGAATTAATGATGAGAGAGATGAATAATTAAGATGCACTTACAT	22219
QY	2221	TTGAAAGAGCAAGTATTATTTCTGCACTTTTGTCCAAATGTAACTGTACTTATATT	2280
Db	22220	TTGAAAGAGCAAGTATTATTTCTGCACTTTTGTCCAAATGTAACTGTACTTATATT	22279
QY	2281	TTATGTATGATGTGAATACCAATACATATATATTTCTAATGTATGATGCAAGTCAAA	2340
Db	22280	TTATGTATGATGTGAATACCAATACATATATATTTCTAATGTATGATGCAAGTCAAA	22339
QY	2341	AATCTTTAAAGCACTGGCCTGTGCTPAACTTCCTTATTTTGACAGAGAGAAATCCAAATC	2400
Db	22340	AATCTTTAAAGCACTGGCCTGTGCTPAACTTCCTTATTTTGACAGAGAGAAATCCAAATC	22399
QY	2401	TGAGAGAGCAACAACATTTTGCCTGAGGTTATAGAACCAAGCTTATGCAATGCTPAAAGTGA	2460
Db	22400	TGAGAGAGCAACAACATTTTGCCTGAGGTTATAGAACCAAGCTTATGCAATGCTPAAAGTGA	22455
QY	2461	TTCTTACTTAAATATCTTTTCCCACTAGTGCAATACGCACTTCTAATCTGTGTGGCTGA	2520
Db	22460	TTCTTACTTAAATATCTTTTCCCACTAGTGCAATACGCACTTCTAATCTGTGTGGCTGA	22519
QY	2521	AATACAGAAATATATATAGTGAAGACAGATACACAAGTCTGGGGAAAAATATTTGGGTAGGTG	2580
Db	22520	AATACAGAAATATATATAGTGAAGACAGATACACAAGTCTGGGGAAAAATATTTGGGTAGGTG	22579
QY	2581	GCTGAGAGCCTCATTTTCTTAAAGAAATGTGAGCCTTAGGCAAGGATGTGTGCTCAACCT	2640
Db	22580	GCTGAGAGCCTCATTTTCTTAAAGAAATGTGAGCCTTAGGCAAGGATGTGTGCTCAACCT	22639
QY	2641	ATATATTCACACATTTTGGAGGCGCAAGTCAAGAAAGATGCGTTGAACCCCAAGAGTTCAAGA	2700
Db	22640	ATATATTCACACATTTTGGAGGCGCAAGTCAAGAAAGATGCGTTGAACCCCAAGAGTTCAAGA	22699

QY 2701 CTACATGGGCAACATAGCAAGACCTCATCTCTACAAAATTTTAAAAATCAGCTGAGCA 2760
DB 22700 CTACATAGGGGCAACATAGCAAGACCTCATCTCTACAAAATTTTAAAAATCAGCTGAGCA 22759
QY 2761 TGGTGGGATACGCTGTAGTCCACCTACCTGGAGAGCTAGGGTGGATCGCTTACA 2820
DB 22760 TGGTGGGATACGCTGTAGTCCACCTACCTGGAGAGCTAGGGTGGATCGCTTACA 22819
QY 2821 CAGGAGTTTGAAGCTTAAGGTAGCCATGATCACAACCTGCACTCCAGCTTGAATGACAG 2880
DB 22820 CAGGAGTTTGAAGCTTAAGGTAGCCATGATCACAACCTGCACTCCAGCTTGAATGACAG 22879
QY 2881 AGGAAAGCCCTGCTCCCTAAAAAGAAATGAGATTTTATTCCTTAACAGCTAAGT 2940
DB 22880 AGGAAAGCCCTGCTCCCTAAAAAGAAATGAGATTTTATTCCTTAACAGCTAAGT 22939
QY 2941 CATTAGTCATTAAAGTTTGAAGTTGAGAGAAATATATGATCAGAAAGAAATTTATCATCT 3000
DB 22940 CATTAGTCATTAAAGTTTGAAGTTGAGAGAAATATATGATCAGAAAGAAATTTATCATCT 22999
QY 3001 GTGGTCTGTAGGATATTTGAAGAAATTAAGAGACTAGAGTCAAGGATTCACCTTAAAGTG 3060
DB 23000 GTGGTCTGTAGGATATTTGAAGAAATTAAGAGACTAGAGTCAAGGATTCACCTTAAAGTG 23059
QY 3061 TTTGTTGTTGTTTGAAGACAGAGTCTTTTGTACCCAGGCTAGAGTGAAGTG 3120
DB 23060 TTTGTTGTTGTTTGAAGACAGAGTCTTTTGTATCCAGGCTAGAGTGAAGTG 23119
QY 3121 CAGTCATGAGTCAACCGAGCCTCAAACTCCAGCCTCAAAATTTATCTTCCAGCTCGGCT 3180
DB 23120 CAGTCATGAGTCAACCGAGCCTCAAACTCCAGCCTCAAAATTTATCTTCCAGCTCGGCT 23179
QY 3181 CCCAAGTGTGGAATTAACAGGTGTAGCCAAAGGTTTATGTAGTGTGCTGCTGAGT 3240
DB 23180 CCCAAGTGTGGAATTAACAGGTGTAGCCAAAGGTTTATGTAGTGTGCTGCTGAGT 23239
QY 3241 GCCTCTCAAACTTCAGTGAGACAGCAAGTACCGGGAACCTGACTCAACAAAGCTGGGTT 3300
DB 23240 GCCTCTCAAACTTCAGTGAGACAGCAAGTACCGGGAACCTGACTCAACAAAGCTGGGTT 23299
QY 3301 TAAAGCTGAGCCTCTGCTATCTTAACTAGTCAAGCTGAGTGAATGCTGTCAAG 3360
DB 23300 TAAAGCTGAGCCTCTGCTATCTTAACTAGTCAAGCTGAGTGAATGCTGTCAAG 23359
QY 3361 GCCAAGCACTGAGACAGAGATCTAGTGAATTAAGTGAATTAAGTGAATTAAGT 3420
DB 23360 GCCAAGCACTGAGACAGAGATCTAGTGAATTAAGTGAATTAAGTGAATTAAGT 23419
QY 3421 AGTGAATTAAGTGAAGAGATGTAATCAATCAAGAAAGAGTTCACCTCTTG 3480
DB 23420 AGTGAATTAAGTGAAGAGATGTAATCAATCAAGAAAGAGTTCACCTCTTG 23479
QY 3481 GTAAATGTAGCATGAGAGAGAGAGTGGGCGCAACATTAAGTGTGTGTGACT 3540
DB 23480 GTAAATGTAGCATGAGAGAGAGAGTGGGCGCAACATTAAGTGTGTGTGACT 23539
QY 3541 GACGAGAGAAATTTGATGCTCTATTAAACAGAAATAGGAGAGAGTGTGTGAGAGAA 3600
DB 23540 GACGAGAGAAATTTGATGCTCTATTAAACAGAAATAGGAGAGAGTGTGTGAGAGAA 23599
QY 3601 GAGAGTCTCTGTTTGAAGCTGTGAGTCCAGGTGAGACAGAGTCTCAAGAGGAAT 3660
DB 23600 GAGAGTCTCTGTTTGAAGCTGTGAGTCCAGGTGAGACAGAGTCTCAAGAGGAAT 23659
QY 3661 GAGCACTAGGCACTTAAAGAAATCTGTGCTCAAGAGAGCTGTGAGCTGAGCTGT 3720
DB 23660 GAGCACTAGGCACTTAAAGAAATCTGTGCTCAAGAGAGCTGTGAGCTGAGCTGT 23719
QY 3721 AGATCTGAGGGTCAATGAGCATAGAGTTTAAAGAGCAAGAGTAGGCAACCAAGAGAGC 3780
DB 23720 AGATCTGAGGGTCAATGAGCATAGAGTTTAAAGAGCAAGAGTAGGCAACCAAGAGAGC 23779

QY 3781 AANTACAAAGAGAGAGAGCTGATGATGAGACTTTGCTTTTATGATGAGAGAGCA 3840
DB 23780 AANTACAAAGAGAGAGAGAGCTGATGATGAGACTTTGCTTTTATGATGAGAGAGCA 23839
QY 3841 ACAGAAATGAGAGATGAAAGGAGACAGCTGTAGGAATGTAGACATCTGAAAAAAA 3900
DB 23840 ACAGAAATGAGAGATGAAAGGAGACAGCTGTAGGAATGTAGACATCTGAAAAAAA 23899
QY 3901 ATACACACTGTCAATGAGATCAAGGAGAGAAATTTCAAGAAAGAGGATATGTTGACA 3960
DB 23900 ATACACACTGTCAATGAGATCAAGGAGAGAAATTTCAAGAAAGAGGATATGTTGACA 23959
QY 3961 GATTTAACAAGCATCAGAGATTAACAGCTAAAGATCATCTTACATGCAATGACCTTGTGG 4020
DB 23960 GATTTAACAAGCATCAGAGATTAACAGCTAAAGATCATCTTACATGCAATGACCTTGTGG 24019
QY 4021 ATTTGAGAGGACACACTAATAAATTAAGAAATTTATGTGGATATATGAGGACAAAG 4080
DB 24020 ATTTGAGAGGACACACTAATAAATTAAGAAATTTATGTGGATATATGAGGACAAAG 24079
QY 4081 GAAAGTTATTCAAATCAAGCAAGTGGAGTGGAGTGGATGAGTTCCAAAGTGAAGCA 4140
DB 24080 GAAAGTTATTCAAATCAAGCAAGTGGAGTGGAGTGGAGTGGATGAGTTCCAAAGTGAAGCA 24139
QY 4141 TCAGTGAATGTGGAAAGGGGCAACAGACATCCATGCCATCCAGGCAAGCCCTCCA 4200
DB 24140 TCAGTGAATGTGGAAAGGGGCAACAGACATCCATGCCATCCAGGCAAGCCCTCCA 24199
QY 4201 GAAGCTCCATGAGAGTTTCAAGTATCCAGAAAGTCTGTGATCCTTAATCTTTCTGGGTT 4260
DB 24200 GAAGCTCCATGAGAGTTTCAAGTATCCAGAAAGTCTGTGATCCTTAATCTTTCTGGGTT 24259
QY 4261 TGCATAGGCTTCAATGTGTAGGCAATGATTTATTAACATATGAGGCACTGTGATCACTT 4320
DB 24260 TGCATAGGCTTCAATGTGTAGGCAATGATTTATTAACATATGAGGCACTGTGATCACTT 24319
QY 4321 AACCTTCAACCCCTCTCCCTCCCTAATCATATCTGTGGCTTCCAGAGCAACAGTCCCTA 4380
DB 24320 AACCTTCAACCCCTCTCCCTCCCTCAATCATATCTGTGGCTTCCAGAGCAACAGTCCCTA 24379
QY 4381 TCCCTAAGCTAACCAATGTGCTGCAAGCTATCAGTCAACTTCAAAAAAGACATCACTTGG 4440
DB 24380 TCCCTAAGCTAACCAATGTGCTGCAAGCTATCAGTCAACTTCAAAAAAGACATCACTTGG 24439
QY 4441 GAGATTCCTAAGATTTTGAAGATTTGCTGCTGAGAAATTAAGTGAAGATCAATATATAT 4500
DB 24440 GAGATTCCTAAGATTTTGAAGATTTGCTGCTGAGAAATTAAGTGAAGATCAATATATAT 24499
QY 4501 TTCACAATATCAACATGCTGCTAATTTATCAAGGCGCATTAATGTTTAAACAAAG 4560
DB 24500 TTCACAATATCAACATGCTGCTAATTTATCAAGGCGCATTAATGTTTAAACAAAG 24559
QY 4561 AGGTATTAATTCAGATTTTCTTTTATTAAGCTTACATGATGACAGTGTGTAATAG 4620
DB 24560 AGGTATTAATTCAGATTTTCTTTTATTAAGCTTACATGATGACAGTGTGTAATAG 24619
QY 4621 ATTTGGAGTGAAGGCAATCTTTTGAATGTATATTCCTGACCCCACTTCTC 4680
DB 24620 ATTTGGAGTGAAGGCAATCTTTTGAATGTATATTCCTGACCCCACTTCTC 24679
QY 4681 CTGTGTTTCTTACCTCTCTCCCTACTCACAGAAAACTTCTCCTCCTCAT 4740
DB 24680 CTGTGTTTCTTACCTCTCTCCCTACTCACAGAAAACTTCTCCTCCTCAT 24739
QY 4741 TCCCTGAATGCTGTGCTGTTAAGGTTCCAGCTTGAACAGGCTTAATCAAGAACAC 4800
DB 24740 TCCCTGAATGCTGTGCTGTTAAGGTTCCAGCTTGAACAGGCTTAATCAAGAACAC 24799
QY 4801 AGTGTATCAGATGTGAGATGATGCTGGGAGAAAGTGAACATATGAGACCAATTAACCTTA 4860
DB 24800 AGTGTATCAGATGTGAGATGATGCTGGGAGAAAGTGAACATATGAGACCAATTAACCTTA 24859
QY 4861 GCCGAACTGACGGGAAAAACAAGAGTCAAGATATTTTCTGATCTGATATTAATA 4920

```

Db      24860 GCGGAACTGACGGGAAAAACAAGAGTCAAGCATATTTTCTGAGTCAATTTAA 24919
Qy      4921  ATGATGATGATGTCATTCATCTGTGATAGATCAAGAAAGAAAAATTTATTTGGAGAG 4980
Db      24920 ATGGATGATGATGTCATTCATCTGTGATAGATCAAGAAAGAAAAATTTATTTGGAGAG 24979
Qy      4981  ATACCATGAATGTTGTTTATGACATGCTAAGTTGAGGTGATTTATGGATGTCAGGCCA 5040
Db      24980 ATACCATGAATGTTGTTTATGACATGCTAAGTTGAGGTGATTTATGGAGGTCA 25039
Qy      5041  GCTCCGACATGTTGGGCTTAAAGTAAAGGCAATCTGATGGATGATTAAGATTTTA 5100
Db      25040 GCTCCGACATGTTGGGCTTAAAGTAAAGGCAATCTGATGGATGATTAAGATTTTA 25099
Qy      5101  AATCATCAGAAATACGTTGTTTCTTATGAGCATCTGATGGGTAGATGCTAAGGAGC 5160
Db      25100 AATCATCAGAAATACGTTGTTTCTTATGAGCATCTGATGGGTAGATGCTAAGGAGC 25159
Qy      5161  ATGTGTAGATGATTAACAAGATCAAAGACGGAACCTTAAAGATTAACAATATGTTATTA 5220
Db      25160 ATGTGTAGATGATTAACAAGATCAAAGACGGAACCTTAAAGATTAACAATATGTTATTA 25219
Qy      5221  TTTATATATTTTATTAATGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 5280
Db      25220 TTTATATATTTTATTAATGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 25279
Qy      5281  ACGGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5340
Db      25280 ACGGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25339
Qy      5341  ACCTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5400
Db      25340 ACCTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25399
Qy      5401  CAGGACCCACCAACCTCACTGCTAATTTTGTATTTTATTTTATTTTATTTTATTTTATTT 5460
Db      25400 CAGGACCCACCAACCTCACTGCTAATTTTGTATTTTATTTTATTTTATTTTATTTTATTT 25459
Qy      5461  ATGTTGGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5520
Db      25460 ATGTTGGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25519
Qy      5521  AAAAGTCTGGGATTAACAGTATGAGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5580
Db      25520 AAAAGTCTGGGATTAACAGTATGAGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25579
Qy      5581  ATGGGCTCTTGCATGTTGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5640
Db      25580 ATGGGCTCTTGCATGTTGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 25639
Qy      5641  GCTTGGCTCTCAAAAGTCTGGGATTAACAGTATGAGTCCCTGCTGCTGCTGCTGCTGCTGCT 5700
Db      25640 GCTTGGCTCTCAAAAGTCTGGGATTAACAGTATGAGTCCCTGCTGCTGCTGCTGCTGCTGCT 25699
Qy      5701  CAATATATTAGATTTTATTTAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 5760
Db      25700 CAATATATTAGATTTTATTTAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 25759
Qy      5761  AAAAAATTAGAGATTTTGGAAACCAAAAGAAAGAGTCCCTGAGAAAGAAAGAGATGCTCT 5820
Db      25760 AAAAAATTAGAGATTTTGGAAACCAAAAGAAAGAGTCCCTGAGAAAGAAAGAGATGCTCT 25819
Qy      5821  ATGATGCCAAATGCTCCAAAGATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAA 5880
Db      25820 ATGATGCCAAATGCTCCAAAGATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAA 25879
Qy      5881  GTCATGGGAAACCATGCTAAAGAAACATTTGTGATGACACAAATGTTGCAAAAGCATTTT 5940
Db      25880 GTCATGGGAAACCATGCTAAAGAAACATTTGTGATGACACAAATGTTGCAAAAGCATTTT 25939
Qy      5941  TATAGGGGATGAATTTTGTATTTTCAAGAGCAAAAGTCCCTAACAATGGCAAGATCTTA 6000

```

```

Db      25940 TATAGGGGATGAATTTTGTATTTTCAAGAGCAAAAGTTCATACATGGCAAGATCTTA 25999
Qy      6001  GTGTGTGACCCACGGAGTTAGTGTCTGAAGTGGATTTGAAAGCAGATCATTTGGAGCTGA 6060
Db      26000 GTGTGTGACCCACGGAGTTAGTGTCTGAAGTGGATTTGAAAGCAGATCATTTGGAGCTGA 26059
Qy      6061  GGTGGCTAGCTAGTCTTCAAGTACAGACATTAATGTCTAGAGTCAACAGCTGTGATCAAG 6120
Db      26060 GGTGGCTAGCTAGTCTTCAAGTACAGACATTAATGTCTAGAGTCAACAGCTGTGATCAAG 26119
Qy      6121  TGCCCATCTTCAAGTAAAGACAGAGAGATTTGAGATTCAGTAAATGACCGCTTAA 6180
Db      26120 TGCCCATCTTCAAGTAAAGACAGAGAGATTTGAGATTCAGTAAATGACCGCTTAA 26179
Qy      6181  GAAAGTAAATGAAAGATGTGG 6201
Db      26180 GAAAGTAAATGAAAGATGTGG 26200

RESULT 2
ADJ84182/c
ID      ADJ84182 standard; DNA; 76698 BP.
XX
AC      ADJ84182;
XX
DT      06-MAY-2004 (first entry)
XX
DE      Human c-rsf genomic DNA which is a target for antisense therapy.
XX
KW      antimicrobial; antiinflammatory; cytostatic; infection; inflammation;
KW      tumour formation; antisense therapy; human; c-rsf; ds; target.
XX
OS      Homo sapiens.
XX
PN      MO2004003134-A2.
XX
PD      08-JUN-2004.
XX
PF      12-JUN-2003; 2003MO-US018481.
XX
PR      26-JUN-2002; 2002US-0392020P.
XX
PA      (ISIS-) ISIS PHARM INC.
XX
PI      Monica BP, Freiler SM, Manoharan M, Gaarde WA;
XX
WPI; 2004-083026/08.
XX
PT      Decreasing levels of a preselected cellular mRNA in a cell, useful for
PT      treating or preventing a disease associated with a preselected cellular
PT      mRNA, comprises binding to the mRNA an antisense compound targeted to a
PT      splice site on the mRNA.
XX
PS      Example 14; SEQ ID NO 1; 174bp; English.
XX
CC      The invention relates to a novel method for decreasing levels of a
CC      preselected cellular mRNA in a cell or tissue comprising binding to the
CC      preselected cellular mRNA an antisense compound which is specifically
CC      hybridisable with a splice site on the mRNA and which is not a substrate
CC      for RNase H when bound to the RNA. The method of the invention has
CC      antimicrobial, antiinflammatory and cytostatic applications and may be
CC      useful for decreasing levels of a preselected cellular mRNA in a cell or
CC      tissue and thus for treating or preventing a disease or condition
CC      associated with a preselected cellular mRNA or with a preselected target
CC      cellular protein, particularly infection, inflammation or tumour
CC      formation. The current sequence is that of the human c-rsf genomic DNA of
CC      the invention which is a target for antisense therapy.
XX
SQ      Sequence 76698 BP; 21036 A; 16662 C; 17183 G; 21716 T; 0 U; 101 Other;
Query Match      6.8%; Score 419.8; DB 12; Length 76698;
Best Local Similarity 52.3%; Pred. No. 7.9e-65;
Matches 1329; Conservative 0; Mismatches 1152; Indels 58; Gaps 16;

```


Db 3561 TCTGAATGCTTGAAGAGAAATAGACATGCCCATGATGTCAGATCAGAGGAGAAAT 3502
 Qy 2616 AGGACAGGATGATGATGCTCACAACCTAATTCAGACATTTGGAGGAGCAAGTCAAGAA 2675
 Db 3501 AGGCGGGGCGAGTGGCTCAGCCTTTATCCCAACATTTGGAGGCTGAGGAGGAGCA 3442
 Qy 2676 ATGCGTTGAACCCAGATTCAAGACTAGCAGTGGGCAACATAGCAAGACCTCATCTTAC 2735
 Db 3441 ATCATCTGAGTCCAGAGATTAAAGACACCTGGCAACATGACAAATCCCATCTCTAC 3382
 Qy 2736 AAAAAATTTAAATAGCTGAGCAGATGAGTACGCTGATGCTCCACCTTACCTGGGA 2795
 Db 3381 TAAATAATACAAATAATAGCTGGGATGATGCGCATGCTTGTAGTCCAGCTACTGTGA 3322
 Qy 2796 AGCTAGTGGGTGGA--TCGCTTGACACAGAGATTGAGGCTAAGGTGAGCCATGATCAC 2853
 Db 3321 GGCTGAGGACAAAGATTGCTTGAACCGGGGAGTTGAGGTTGCACTGAGCCAAAGTTGT 3262
 Qy 2854 ACACTGCACTCCAGCTTGAATGACAGAGAGAACCTTGTCCCTAATAAAGAAAGAAATG 2913
 Db 3261 GCCACTGCACTCCAGCTGGGCGACAGAGGAGAACCTGTCTCAAAAAA 3202
 Qy 2914 TGGATTTTATTCCTTAGAC 2932
 Db 3201 AAAAGATGTACCGTTGGC 3183

RESULT 3

AAT15610
 ID AAT15610 standard; cDNA; 1930 BP.

AC AAT15610;

DT 25-JUN-1996 (first entry)

DE CROC-1 cDNA encodes c-fos promoter activating protein.

XX CROC-1; CROC-4; c-fos promoter activating protein; signal transducer;
 KM polyomavirus large T antigen; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..516

FT /tag= a

FT /product= "CROC-1"

XX W09601899-A1.

XX 25-JAN-1996.

XX 05-JUL-1995; 95WO-US007874.

XX 08-JUL-1994; 94US-00272412.

XX (SCHE) SCHERING CORP.

XX Lin SL, Rothofsky ML;

XX MPI, 1996-097629/10.

XX P-PSDB; AAR90769.

XX Mammalian cell lines contg. recombinant vectors encoding c-fos promoter

XX PT activating proteins - used to identify nucleic acids encoding signal

XX transducing molecules which activate promoters.

XX Claim 13; Page 29-31; 41pp; English.

XX CROC-1 cDNA encodes a c-fos promoter activating protein of approx. 19 kd

XX with an acidic amino terminal half and a basic carboxy terminus. The

XX protein includes a kinase target domain which contains phosphorylation

XX sites for a variety of kinases involved in signal transduction. CROC-1

CC mRNA is approx. 2.3 kb in length and present in all tissues examined. The
 CC cDNA is used in vectors, operatively linked to a nucleic acid encoding a
 CC polyomavirus large T antigen. Mammalian cell lines contg. these vectors,
 CC and vectors comprising a polyoma origin of replication and a nucleic acid
 CC suspected to encode and activating protein of the promoter, can be used
 CC to identify signal transducing molecules

XX Sequence 1930 BP; 524 A; 487 C; 404 G; 515 T; 0 U; 0 Other;

Query Match 6.5%; Score 402.6; DB 2; Length 1930;

Best Local Similarity 92.5%; Pred. No. 6.5e-62;

Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

Qy 16 CAATTTATGAAGCCAAATATACAGCTTAAATATGATGAGTGAATATGATGATGATG 75
 Db 245 CAATTTATGAAGCCAAATATACAGCTTAAATATGATGAGTGAATATGATGATGATG 304
 Qy 76 AACTCCCCCTTGTATGATTTGATGACAAATAATATATGATGATGATGATGATGATG 135
 Db 305 CACCCCTTGTATGATTTGATGACAAATAATATATGATGATGATGATGATGATGATG 364
 Qy 136 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 195
 Db 365 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 424
 Qy 196 AAGTTATCTTCAAGAGCTTACAGCGCTTAATGATGCTTAAGAAATGTAAGACGCCCTC 255
 Db 425 AAGTTATCTTCAAGAGCTTACAGCGCTTAATGATGCTTAAGAAATGTAAGACGCCCTC 484
 Qy 256 AGCCATGTAAGAGAGCTTACAGCAATGATGATGATGATGATGATGATGATGATGATGATG 315
 Db 485 AGCCGCGGAGAGAGCTTACAGCAATGATGATGATGATGATGATGATGATGATGATGATG 544
 Qy 316 CTTCCTCCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 375
 Db 545 CTTCCTCCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 604
 Qy 376 AGCTTGTAGAGCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 435
 Db 605 GCTTGTAGAGCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 664

RESULT 4

AAV16738
 ID AAV16738 standard; cDNA; 1930 BP.

AC AAV16738;

DT 25-MAR-2003 (revised)

DT 18-JUN-1998 (first entry)

XX cDNA encoding human CROC-1 protein.

XX CROC-1; activating protein; promoter; proto-oncogene; c-fos; antagonist;

XX de.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..516

FT /tag= a

FT /product= "CROC-1"

XX US5736331-A.

XX 07-APR-1998.

XX 18-OCT-1995; 95US-00544900.

Db 704 AGCCGCGGAGAGACAGTCTTACAGCAATTAATCAAAAAAGAACACAGGCCCTTCCC 763
 Qy 316 CTTCCCGCATCTGATGATGACAGTCTTCAATTTCCATAGTAGTAATTTTCTAGATAC 375
 Db 764 CTTCCCGCATCTGATGATGACAGTCTTCAATTTCCATAGTAGTAATTTTCTAGATAC 823
 Qy 376 AGCTTGATAGCTCAAAAGTCTGGAAGAAAGAAAGTCCCATCAAGGAATTTATCTTAAG 435
 Db 824 GCTTGTGATGACCTCAAAAGTCTGGAAGAAAGAAAGTCCCATCAAGGAATTTATCTTAAG 883
 Qy 436 ATACTGTAATGATGATCTAA-TTTTGTACATTTGGAATATATAGTTGT 483
 Db 884 ATACTGTAATGATGATCTAA-TTTTGTACATTTGGAATATATAGTTGT 932

RESULT 6
 ADL62300
 ID ADL62300 standard; DNA; 3802 BP.

XX ADL62300;
 AC 20-MAY-2004 (first entry)
 DT
 XX Human ovarian cancer DNA marker #20512.
 DE Human ovarian cancer; ds; tumour; cytostatic; DNA marker.
 XX Homo sapiens.
 OS
 XX WO200170979-A2.
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009126.
 XX
 PR 21-MAR-2000; 2000US-0191031P.
 PR 25-MAY-2000; 2000US-0207124P.
 PR 15-JUN-2000; 2000US-0211940P.
 PR 07-JUL-2000; 2000US-0216820P.
 PR 25-JUL-2000; 2000US-0220661P.
 PR 21-DEC-2000; 2000US-0257672P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI Lee J, Lillie J;
 XX
 XX WPI; 2001-611502/70.
 DR
 XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
 PT cancer cells as compared to their normal non-cancerous ovarian cells are
 PT used to characterize stage, grade, histological type of ovarian cancer.
 XX
 PS Disclosure; SEQ ID NO 20512; 106pp; English.
 XX
 XX The invention relates to nucleic acid markers which are overexpressed in
 CC ovarian cancer cells as compared to their expression in normal (i.e. non-
 CC cancerous) ovarian cells. The invention also relates to polypeptides
 CC encoded by the markers, antibodies that selectively bind to the
 CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
 CC of developing ovarian cancer involving inhibiting expression of a gene
 CC corresponding to a marker of the invention and a method of treating a
 CC patient afflicted with ovarian cancer comprising providing to cells of
 CC the patient an antisense oligonucleotide complementary to a marker of the
 CC invention. The markers are useful for assessing if a patient is afflicted
 CC with ovarian cancer, which involves comparing the level of expression of
 CC a marker in a patient sample and a normal level of expression of the
 CC marker in a control non-ovarian cancer sample. A difference between the
 CC expression levels indicates ovarian cancer. The level of expression of a
 CC marker corresponds to a secreted protein or to a transcribed
 CC polynucleotide or its portion. The level of expression of the marker is
 CC assessed by detecting the presence in the sample, a protein or protein
 CC fragment corresponding to the marker. The presence of protein or protein
 CC fragment is detected using an antibody that specifically binds with the

CC protein or protein fragment. Alternatively, the level of expression of
 CC the marker is assessed by detecting the presence of a transcribed
 CC polynucleotide which anneals with the marker or anneals with a portion of
 CC the polynucleotide comprising the marker, under stringent conditions. The
 CC marker is also used for monitoring the progression of ovarian cancer in a
 CC patient which involves detecting expression of the marker in a patient
 CC sample at a first point in time, repeating the method at a subsequent
 CC time and comparing the level of expression. The method is carried out
 CC using an ovarian tissue sample. A composition comprising a marker,
 CC polypeptide or antibody of the invention is used to treat ovarian cancer.
 CC This sequence represents a human ovarian cancer DNA marker of the
 CC invention.
 XX
 XX Sequence 3802 BP; 979 A; 967 C; 837 G; 1019 T; 0 U; 0 Other;

Query Match 6.5%; Score 402.6; DB 5; Length 3802;
 Best Local Similarity 92.5%; Pred. No. 6.9e-62;
 Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

Qy 16 CAATTTAGAAAGCCAAATATACAGCTTAAATATGATGTGACCTTAATACCCAGAG 75
 Db 438 CAATTTAGAAAGCCAAATATACAGCTTAAATATGATGTGACCTTAATACCCAGAG 497
 Qy 76 AACGCCCTTGTAGATTTGTGAACAAATTAATATGAGTAGAGTTAATAGTTCTAATG 135
 Db 498 CACCCCTTGTGTAGATTTGTGAACAAATTAATATGAGTAGAGTTAATAGTTCTAATG 557
 Qy 136 GAATGTGAACCCAGAGCCATATCAGCGCTAGCAAAATGGCAGAAATTCATATATCATCA 195
 Db 558 GAGTGTGAGCCAGAGCCATATCAGCGCTAGCAAAATGGCAGAAATTCATATATCATCA 617
 Qy 196 AAGTTATCTTCAAGAGCTTCAAGCGCTTAATGATGTCTTAAGAAATGTGAAACGCCCTC 255
 Db 618 AAGTTGTCTTCAAGAGCTTCAAGCGCTTAATGATGTCTTAAGAAATGTGAAACGCCCTC 677
 Qy 256 AGCCATCTGAAGGAGAGTGTACAGCAATGTGATCAAAAAAGAAACACAGGCCCTTCCC 315
 Db 678 AGCCGCGGAGAGAGTGTACAGCAATGTGATCAAAAAAGAAACACAGGCCCTTCCC 737
 Qy 316 CTTCCCGCATCTGATGATGACAGTCTTCAATTTCCATAGTAGTAATTTTCTAGATAC 375
 Db 738 CTTCCCGCATCTGATGATGACAGTCTTCAATTTCCATAGTAGTAATTTTCTAGATAC 797
 Qy 376 AGCTTGATAGCTCAAAAGTCTGGAAGAAAGAAAGTCCCATCAAGGAATTTATCTTAAG 435
 Db 798 GCTTGTGATGACCTCAAAAGTCTGGAAGAAAGAAAGTCCCATCAAGGAATTTATCTTAAG 857
 Qy 436 ATACTGTAATGATGATCTAA-TTTTGTACATTTGGAATATATAGTTGT 483
 Db 858 ATACTGTAATGATGATCTAA-TTTTGTACATTTGGAATATATAGTTGT 906

RESULT 7
 AAS63021
 ID AAS63021 standard; cDNA; 2144 BP.
 XX
 XX AAS63021;
 AC 29-JAN-2002 (first entry)
 DT
 XX Cell death protective sequence CNT-00726.
 DE
 XX Human; protective sequence; cell death; cerebral oedema; infection;
 KW meningitis; degenerative disease; Alzheimer's disease; heart disease;
 KW motor neuron disease; demyelinating disease; multiple sclerosis; asthma;
 KW nutritional condition; peripheral nervous system disorder; ischaemia;
 KW diabetic neuropathy; autoimmune haemolytic anaemia; respiratory system;
 KW oral cavity; gastrointestinal tract; liver; cirrhosis; pancreatitis;
 KW polycystic renal disease; urinary tract; genitalia; endometriosis;
 KW breast; chronic mastitis; thyroid gland; Hashimoto's thyroiditis;
 KW adrenal gland; skin; psoriasis; muscular atrophy; bone marrow;
 KW osteoporosis; cancer; autoimmune disease; ss.

OS Homo sapiens.
 XX MO200176532-A2.
 XX 18-OCT-2001.
 PD 09-APR-2001; 2001WO-US011655.
 PF 11-APR-2000; 2000US-00547596.
 XX (COGE-) COGENT NEUROSCIENCE INC.
 PA Barney S, Thomas MB, Portbury SD, Puranam K, Katz LC, Lo DC;
 PI WPI; 2002-017408/02.
 DR Novel nucleic acids referred as protective sequences and their encoded
 XX products for diagnosing, treating diseases involving cell death,
 PT including neurological disorders e.g. stroke and for identifying
 PT modulators.
 PS Claim 2; Fig ID; 256p; English.
 XX The invention relates to isolated protective sequence polypeptides (I)
 CC and polynucleotides (II). (II) is useful for transferring a protective
 CC sequence into a cell, which delays and/or prevents the cell from
 CC undergoing cell death. Protective sequences, their products or antibodies
 CC are useful diagnostically, prophylactically, therapeutically or as
 CC targets for treatment and diagnosis of conditions, disorders or diseases
 CC involving cell death. The protective sequences and their products are
 CC useful for preventing or treating disorders of the central nervous system
 CC including neurological and psychiatric conditions, cerebral oedema,
 CC infections such as meningitis, degenerative diseases such as Alzheimer's
 CC and motor neuron disease, demyelinating diseases such as multiple
 CC sclerosis, nutritional conditions, disorders of the peripheral nervous
 CC system including diabetic neuropathy, disorders which cause cell death in
 CC organ systems including blood vessels, heart (ischaemia), blood cells
 CC (autoimmune haemolytic anaemia), respiratory system (asthma), oral
 CC cavity, gastrointestinal tract, liver (cirrhosis), pancreatitis,
 CC poly cystic renal disease, urinary tract, genitalia (congenital
 CC anomalies), endometriosis, breast (chronic mastitis), thyroid gland
 CC (Hashimoto's thyroiditis), adrenal gland, skin (psoriasis),
 CC musculoskeletal system (muscular atrophy), bone marrow or bone
 CC (osteoporosis). The compositions promote cell death and are useful for
 CC treating and/or ameliorating cancer and autoimmune diseases. The
 CC compounds are further useful for treating physiological impacts on organs
 CC caused by infection which induce cell death. (I) is useful to raise an
 CC immune response, as a reagent in assays designed to quantitatively
 CC determine levels of the protein in biological fluids, as markers for
 CC tissues in which the corresponding protein is expressed and to isolate
 CC receptors or ligands. AA62967-AA63142 represent the protective
 CC polynucleotide sequences as described in the invention
 XX
 XX Sequence 2144 BP; 608 A; 511 C; 434 G; 591 T; 0 U; 0 Other;
 SQ
 Query Match 6.5%; Score 401.6; DB 6; Length 2144;
 Best Local Similarity 92.3%; Pred. No. 9.9e-62;
 Matches 434; Conservative 0; Mismatches 34; Indels 2; Gaps 1;
 QY 16 CAATTTATGAAAGCCAAATATACAGCCTTAATAATAGTGGACCTTAATATCCAGAAG 75
 DB 195 CAATTATATGAAACCGAATATACAGCCTTAATAATAGTGGACCTTAATATCCAGAAG 254
 QY 76 AACTCCCTTTGTAAGATTGTAACAAATTAATATATGAGTGAAGTTAATAGTTCTAATG 135
 DB 255 GACCCCTTTGTAAGATTGTAACAAATTAATATATGAGTGAAGTTAATAGTTCTAATG 314
 QY 136 GAATGTTGAACCCGAAGCCATATGACGCTAGCAAAATGGGCAATTAATATATATATCA 195
 DB 315 GAGTGTGACACCAAGACCATATGCTGCTACCAAAATGGGCAAAATTAATATATATATCA 374
 QY 196 AAGTATCTTCAGAGGCTTCAGCGCTAATGATGTCTAAAGAAATATGAAACGCCCTC 255

DB 375 AAGTTGTCTCGAAGAGCTTCGGCGCTAATGATGTCTAAAGAAATATGAAACTCCCTC 434
 QY 256 AGCCATCTGAGAGACAGTGTTCACAGCAATTTGATCAAAAAAGAAAACACAGGCCCTTCCC 315
 DB 435 AGCCGCCGGAAGAGAGTGTTCACAGCAATTTATCAAAAAAGAAAACACAGGCCCTTCCC 494
 QY 316 CTTTCCCCCATCTGATGATGTAAGACAGTCTTCATTTTCATATGATGTAATTTTCTAGATAC 375
 DB 495 CTTTCCCCCATCTGATGATGTAAGACAGTCTTCATTTTCATATGATGTAATTTTCTAGATAC 554
 QY 376 AGCTGTAGAGCTCAAGACTGGAAGAAAGAGCTCCCATTCGAAGAGAAATTTATCTTAAG 435
 DB 555 GCTTGTAGACCTCAAGAGTACCGAAGAGAGCTCCCATTCGAAGAGAAATTTATCTTAAG 614
 QY 436 ATACTGTAAATGATATCTAATTTT--GTACATTGTGAATATATTAAGTTGT 483
 DB 615 ATACTGTAAATGATATCTAATTTTTCGTCATTTGAAATATATTAAGTTGT 664
 RESULT 8
 ABK84356
 ID ABK84356 standard; cDNA; 3335 BP.
 XX
 AC ABK84356;
 XX
 DT 14-AUG-2002 (first entry)
 XX
 DE Human cDNA differentially expressed in granulocytic cells #927.
 XX
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW fungal infection; parasitic infection; protozoal infection;
 KW rheumatoid arthritis; sterile inflammatory disease; psoriasis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 XX
 OS Homo sapiens.
 XX
 PN WO200228999-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 03-OCT-2001; 2001WO-US030821.
 XX
 PR 03-OCT-2000; 2000US-0237189P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;
 XX
 DR WPI; 2002-435328/46.
 XX
 PT Detecting granulocyte activation by detecting differential expression of
 PT genes associated with granulocyte activation, which serves as diagnostic
 PT markers that is useful for monitoring disease states and drug toxicity.
 PS
 PS Claim 1; SEQ ID NO 927; 114p; English.
 XX
 XX The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing the
 CC expression level to an expression level in an unactivated GC, where
 CC differential expression of Gs is indicative of GCA. Also included are
 CC modulating (M2) GA by contacting GC with an agent that alters the
 CC expression of at least one gene in Gs; (2) screening (M3) for an agent
 CC capable of modulating GCA or an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease using the gene expression
 CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease, by detecting the level of

CC expression in a sample of the tissue of gene(s) from G6, where the level
 CC of expression of the gene is indicative of inflammation; (4) treating
 CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease, by contacting a tissue having inflammation with an
 CC agent that modulates the expression of gene(s) from G6 in the tissue. M1
 CC is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful
 CC for screening an agent capable of modulating GCA preferably in an
 CC inflammation in a tissue; M4 is useful for detecting an inflammation
 CC (especially chronic) in a tissue, an allergic response in a subject,
 CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
 CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
 CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
 CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, periodontal disease, also bacterial
 CC infection, viral infection, parasitic infection, protozoal infection,
 CC fungal infection and M5 is useful for treating one of the above
 CC conditions. The present sequence represents a gene differentially
 CC expressed in granulocytes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 3335 BP; 839 A; 857 C; 715 G; 924 T; 0 U; 0 Other;

Query Match 5.9%; Score 365.8; DB 6; Length 3335;

Best Local Similarity 92.3%; Pred. No. 2.2e-55; Matches 396; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

OY 56 TGGAGCTTAATATCCCGAAGAACTCCCTTTGTAAGTTTGTAAACAAAATTATATATGAG 115
 DB 1 TGGAGCTTAATATCCCGAAGAACTCCCTTTGTAAGTTTGTAAACAAAATTATATATGAA 60
 OY 116 TAGAGTTAATAGTTCTTAATGGAATGGTGAACCCAGAGCCATATCAAGCTAGCAAAATG 175
 DB 61 TGGAGTTAATAGTTCTTAATGGAATGGTGAACCCAGAGCCATATCAAGCTAGCAAAATG 120
 OY 176 GCAGAAATCATATATCATCAAAAGTTATCTTCAAGAGCTTCAAGCGCTTAATGATGCTAA 235
 DB 121 GCAGAAATCATATATCATCAAAAGTTATCTTCAAGAGCTTCAAGCGCTTAATGATGCTAA 180
 OY 236 AGAAATATGAAAGCGCTTCAAGCATCTAAGAGACAGTGTAAAGCAATTTGATCAAAAAG 295
 DB 181 AGAAATATGAAAGCGCTTCAAGCATCTAAGAGACAGTGTAAAGCAATTTGATCAAAAAG 240
 OY 296 AAAAACAACAGAGCGCTTCCCTCCCTCCCACTTGAATGTAAGAGCTTCAATTTCCATA 355
 DB 241 AAAAACAACAGAGCGCTTCCCTCCCTCCCACTTGAATGTAAGAGCTTCAATTTCCATA 300
 OY 356 GTAGTAATTTTCTAGATACAGCTTGTAGAGCTCAAAAGTACGAGAAAGAAAGCTCCCAT 415
 DB 301 GTAGTAATTTTCTAGATACAGCTTGTAGAGCTCAAAAGTACGAGAAAGAAAGCTCCCAT 360
 OY 416 CAAAGAAATTTTCTTAAGATCTGTAATATGATACAA-TTTTTGTACATTTGGAAATAT 474
 DB 361 CAAAGAAATTTTCTTAAGATCTGTAATATGATACAAATTTTGTCCATTTGAATATAT 420
 OY 475 ATTAAGTTGT 483
 DB 421 ATTAAGTTGT 429

RESULT 9
 ADN16203
 ID ADN16203 standard; DNA; 264965 BP.

XX AC ADN16203;

XX 01-JUL-2004 (first entry)

XX Human sulfatase SULF2 gene.

XX Human; sulfatase; SULF2; enzyme; cytosolic; thrombolytic;

KW antinflammatory; chromosome 20q12-13.2; gene; ds.

XX Homo sapiens.

OS Homo sapiens

FH Key Location/Qualifiers

FT CDS 91625..190625

FT /tag= a "Human SULF2"

FT /product= "Human SULF2"

FT /note= "Contains many introns"

PN WO200401365-A2.

XX 15-APR-2004.

XX 02-OCT-2003; 2003WO-US031442.

XX 03-OCT-2002; 2002US-00265071.

XX (REGC) UNIV CALIFORNIA.

XX (THIO-) THIOS PHARM INC.

XX Rosen S, Hemmerich S, Tomita M;

XX WPI: 2004-330173/30.

XX P-PSDB; ADN16191.

XX New sulfatase polypeptide, useful in diagnosing and treating tumor,

XX inflammation and thrombosis.

XX Example 5; SEQ ID NO 22; 326bp; English.

XX The present sequence is that of the novel human sulfatase SULF2 gene on
 XX chromosome 20q12-13.2. The gene was isolated from a database search using
 XX human SULF2 cDNA ADN16190. It comprises 21 exons. The encoded sulfatase
 XX exhibits glucosamine-6-sulfatase activity, is an endosulfatase, removes a
 XX sulfate group from glycoproteins and/or proteoglycans, is secreted from
 XX eukaryotic cells, acts on extracellular matrix (ECM) components to remove
 XX a sulfate group resulting in release from the ECM of extracellular
 XX differentiation factors and/or growth factors, is secreted in greater
 XX abundance from a cancer cell than from a non-cancerous cell of the same
 XX type, removes 6-sulfate groups from heparin and is pro-angiogenic. mRNA
 XX encoding the sulfatase shows elevated expression in tumours. Novel human
 XX and murine sulfatase polypeptides and polynucleotides of the invention
 XX can be used in diagnostic applications, for therapeutic agent screening
 XX and in the treatment of a variety of disease conditions, including
 XX tumours, inflammation and thrombosis. Methods are provided of modulating
 XX sulfatase activity and of treating disease conditions using such
 XX modulators, particularly by administering inhibitors of the novel
 XX sulfatases.

XX Sequence 264965 BP; 71910 A; 62351 C; 63451 G; 67116 T; 0 U; 137 Other;

Query Match 5.9%; Score 365.2; DB 12; Length 264965;

Best Local Similarity 52.8%; Pred. No. 4.1e-55; Matches 1289; Conservative 0; Mismatches 1043; Indels 108; Gaps 19;

OY 486 GCCTGCGCGGTGCTGCTCAAGCCTGTATCCAGCCCTTTGGAGGCGCAGTGGCGAGAT 545
 DB 218244 GCCAGCGCGGTGCTGCTGCTGCTGTATCCAAACCTTTGGAGGCGCAGTGGCGAGAT 218303
 OY 546 CATGAGCTCAGAGATTGAGACCAAGCTTACCAACATGTTGAAACCCGCTCTTACTAA 605
 DB 218304 CATGAGCTCAGAGATTGAGACCAAGCTTACCAACATGTTGAAACCCGCTCTTACTAA 218363
 OY 606 GATACAAAATTTAGCAGAGTGTGAGGCGCACACTGTAAACCCGCTGCTGAGAGAG 665
 DB 218364 AATTACAAAATTTAGCAGAGTGTGAGGCGCACACTGTAAACCCGCTGCTGAGAGAG 218423
 OY 666 TGAGGCGAGGAATGCTTGAACCCAGAGGCGAGAG-TGACGCGAGCAAAAGATCACACC 724
 DB 218424 TGAGGCGAGGAATGCTTGAACCCAGAGGCGAGAGTGTGAGTGTGAGCGAGATCTCGCC 218483
 OY 725 AATGCACTGTAGCTGTGATGACAGGCGAAGACTCCAACTCAAAAAAATTTAAAAAAGA 784

Db 218484 ACTGCATCCAGCCTGGGCAAGAGCGAGCTCCGCTCAAAAAAAAAAAAAAAAAAAAA 218543
QY 785 AATATGTAAGTTGTGTATAAACAATAATAGCGAGTGAAGCAAGTCTAAAGCTTA 844
Db 218544 AAGGTCTGGCCCATATATTTATCTCTTCTATGACATATTCATTATGAAAAATCTGAGA 218603
QY 845 TGACCATGTTAAGTAATCTGTGGAAACACTATTAAGGAACTTAACCACTTCCTG 904
Db 218604 AATCCAAAGTTGGTTGGAAAACTCAGGGGCTTCATGTGAAGGTAATTTCTGCAAAA 218663
QY 905 G-AAGTAAGGTTTTGAAAAAGAAATGTTTGAGACAAGAGGTTTAAAGAGTGAAGAAAAA 963
Db 218664 GCTACTGGAATTTAAAGGACCTAATATTGAATTAACCTGACGAATAACCAACCACTC 218723
QY 964 AATTAATAATACAGTTTACGTGTGTGAGAAATGGGATGGAGCTTAATAGAAATCAA 1023
Db 218724 CAATTAATCTGTTTAAATTCAGAAATTTAATGTGACCAAAAAATCTTATTAATAGCAGG 218783
QY 1024 AATGGAATGTTTATGTATGTATAAGAACCTGTGTAAGGTTGAAGACATTACATTATCT 1083
Db 218784 GGAAGCTAGGAAAAAGGAAATTAACGATTCGGGAAAGATTAATACCTGTGACCTCGT 218843
QY 1084 GCAACCATCGGGGACTTTTTTTTATGTATGTCTGGCAATTTAAATAGAG--AGCAG 1141
Db 218844 CCAAAATAGACATGATTTGTCTCAACACACTCTCCTGAGTTTAGCCCTGGGCCACATG 218903
QY 1142 AGAATGTAACAAGTTGATGATGATGATGAAATTCCTGCAACAT-GTGAAGGAAG 1200
Db 218904 ACAAATACAAAAAGATGTCATTAATAAAGTTTATTAATGAAAACATCAGGACAGAAAG 218963
QY 1201 AGACAGTAGGCAAGAGATTGAAGAGATTATCAAGACAGAAATTAATG-TGCTGGCCAG 1259
Db 218964 AATAAATATCCCAAGAGCTAAAGTAAGTATCTCAGTTTAAATGATGTTAATCAC 219023
QY 1260 TGGCATCTAGTCTGAAGTCTAATCTGAGGAGAGAAAGTAAGCAGCTTCTGATAG 1319
Db 219024 ACAAGCATTTGATCTTTCTGTGGCAATTTTTTCTCCATCAGTCTGTTTTCTTT 219083
QY 1320 TTATGAAGAGTGAAGGCTTCAAGACCTAAGGTTGATTAATTAAGAAATGATTT 1379
Db 219084 CATGCTGTGTAATTTCTAACCTTTTATTAATTAACATGATATCATTTTATTAATCAGATGAG 219143
QY 1380 GGAGAAAGAAATACGTGAGAGAGTGAGATTTCAAGCTTGAAGTCTCAATACAG 1439
Db 219144 TAATTTTCAA-----AGATGCTACTTCTGTAATTTGTTTAAAGATTTT 219187
QY 1440 ACACTGTCTAAATGCTTCAAGACATGATCCCTGCCCTCAAGGACCTTACAGCCAAAA 1499
Db 219188 ATTTAATCTTCAGTAGCAATATGTAACAATGAGCTATTCATAGAAATTAACACTGC 219247
QY 1500 CAAGAGATTAAGAAATACACCAACTATTAATTAAGACACTTGTGTAATTAAGAA 1559
Db 219248 CAATCATTAAGAGGTGCAAGAAAGACT---CACTCTAAAGATTAATAATATATA 219303
QY 1560 GAATATGATCTAGTACTGTAGATGTGCAACGGCATCAAGATATCTTCTAGTTTCAAG 1619
Db 219304 CAATTCGTCTACTATGTTTATCTGAGGGAGTTAATACAGAAATATGTTCAAAATATA 219363
QY 1620 AGTTTCAAGTCGGCGGGCGCGGTGCTCAAGCTGTAAATCCAGACATTTGGGAGCCG 1679
Db 219364 AGTAACATA-CAGCCGGGCAATGTGCTCATGCTGTAAATCCCAACATTTGGGAGGCCG 219422
QY 1680 AGCGGGGTGATAC--AGGTCAAGAGATCAAGACATCTCTGTAAACAAGGTGAAC 1737
Db 219423 AGGCAAGCGGATTAACCTGAGTCAAGACTTTGAGCAAGCTGACCAAACTGTGTAAC 219482
QY 1738 CCGTCTCTACAAAAAATATATAAATTAGCAGAGCGGTGAGGGCGCTGTAGTCCCA 1797
Db 219483 CTGTCTCTACTAAAAATGTAATAAATTAGCAGGCGGTGTGATGGCGCTGTAAATTTCA 219542
QY 1798 GCTACTCAGAGGCTAGGCAAGAAATGCGGTGAACCCGGGAGGTAGATTTGC-GTGA 1856
|||||

Db 219543 GCTACTGGGAGGCTGAGCAGAGAAATCATTTGAACCCAGAGGCAAGGTTTGCAGTGA 219602
QY 1857 GCCAGATGCGGCCCACTGGCTCCAGCCTGGG--CGACAGATGAGACTGCTTCAAAA 1914
Db 219603 GCTGAGATCATGCAATTCAGTCTCAGCCTGGGCAAGACAGAACAAATCCGCTTCAAAA 219662
QY 1915 AAAAAAAAAAAAAAAAAAGTTTCAGATCTTAAACACACTGCATTTCAACAGTCTAGA 1974
Db 219663 AAAAAAAAAAGTAAC-----TCCATTCAGCGTTAAAC 219693
QY 1975 TAGAGAGCATGTTACAGGAGAGAAAAATGTTTCAGCAAGAGTACAGATGGGAAATA 2034
Db 219694 AAGATTAATCTGTGAATTTTCAAAATTAATCTTTGTCTACTTTAATTTGAAAAATTAATC 219753
QY 2035 GAGGATATGTTCAAGAGAGAGACCCAGAGTCAATGTTTGTAGGTTTGAAGAAACA- 2093
Db 219754 TCAAAAAATCCTGAATGATGTTACTGATGAGCATTAATGATTAATGATGACTGCAAAAATAG 219813
QY 2094 CAGTGTTTGCAATCTCCAGGTTCCATTAGTCCGTTATAGAAATCAATATAGTGTAGCA 2153
Db 219814 CAATATTTCTTTTAAATAATTTCTGCAATTAATGTAATTAATTAATTAATTAACAAACACAC 219873
QY 2154 ACCGCAATTTTAAAAAATGAATTAATGATGAGAGAGAAATAGAAATATTAGCATGCA 2213
Db 219874 ACATACATATGTAATAGTATAGTATGTTTATGTAATACCTACTCCATTTT---TATA 219929
QY 2214 TTACATTTTGAAGACAGATTAATTTTCTGCAACTTTGCTCAATTTACTGTACT 2273
Db 219930 TCACCTGTTTAAAGTCCACATGTAATAATAG-----TGTATACAAAGCA 219975
QY 2274 TATATTTTATGTAATGATGTAATACAGATACATATATATTTCTGTGTAAGCTGA 2333
Db 219976 AAGAAATACAGGCTTAAGTCAAGGACACTAGCTATTAATTTTCTTATTAAGCAATGATCTTG 220035
QY 2334 GTCAAAAAATCTTTAAGACACTGCGCTGATCTTACCTCTTATTTGCAAGAGAAATC 2393
Db 220036 GACAAATACCAATCTCTGTGCAAGCTTGCTTTTATCTGAGATGGGCA--- 220092
QY 2394 CAAGATCTGAGAGGCAAAACATTTTGCCTGAGGTTATGAACCAAGCTTATGCACTTA 2453
Db 220093 -----TATATGCAAAACCTTGGCATCCCATTTACACTTAGATACCATAGAGTA 220144
QY 2454 AAGATGATCTTAAGTAAATCTTTCCACTAGGCAATCTGACTTCACTTCTGTT 2513
Db 220145 ACCCTGAAGTAAATTAATTAATTAATTTGTCACAGGTTTAAATTTATTAATGATCTTAAT 220204
QY 2514 GGCCTGAATACAGATATATTAATGTAACAAGCATACACAAGTCTGGGAAATATATTGG 2573
Db 220205 CAACCAAGATGTCTTAATTTTT-----TATTTCTATGATTAATAAATAG 220249
QY 2574 GTAGGTGGCTGAGACCTCATTTTCTAAGAAATGTGAACCTTAGCAGGTAATGTGCT 2633
Db 220250 CAGACCAAGATTTGAATCTCAGGTTGACGTCAAGATTAATCTCCGCGCAAGACCACTGCT 220309
QY 2634 CACACCTTAATTTCAAGCACTTTGGAGAGCCAAAGTCAAGAAATGATGCTGAACCAAGAG 2693
Db 220310 CACAACTGTAATCCAGCAATTTTGGAGGCTGAGGCAAGCACTACTTGAAGTCAAGAG 220369
QY 2694 TTCAAGACTAGATGGGCAATACCAAGACCTCATCTTCAAAAAATTTAAATTAATCAG 2753
Db 220370 TTGAGAGCAAGCTGGAACAATGTGTGAACCCCATCTCTAATAAATAAATAAATTAG 220429
QY 2754 CTGAGCATGTGTGCAATAGCCTGTATGTCCACCTTACTGGAAGTGA-GGTGGGTGATC 2812
Db 220430 CCGGCAATGTGTGCAAGTGTCTGTATGTCMAACTGTTTGGAGGCGCAAGGCAAGGAAATTC 220489
QY 2813 GCTTG-ACACAGAGATTTGAGGCTAAGGTGAGCCATGATCAACAATCTGCACTCAAGTT 2871
Db 220490 ACTTAACTCAAGAGTGAAGGTGCAATGAGCCGAATCCACACATTTGCACTCAAGCTT 220549
QY 2872 GAGTGACAGAGAAAGACCTGTCTCTTAAAAAAGAAAGAA 2911
Db 220550 GGGCAACAGAGCAAGACTCCGCTCAACAAAAAATAAATAA 220589
|||||

RESULT 10
ID ABS56563 standard; DNA; 268685 BP.
XX
AC ABS56563;
XX
XX 27-JAN-2003 (first entry)
XX
DE Human SULF2 genomic DNA sequence.
XX
KW SULF2; ds; glucosamine-6-sulphatase; cancer; ischaemia; tumour;
KW angiogenesis; coronary; carotid; arterial occlusive disease;
KW peripheral arterial disease; atherosclerosis; myointimal hyperplasia;
KW thromboangiitis obliterans; thrombotic disorder; vasculitis;
KW heart attack; myocardial infarction; vascular death; inflammation;
KW sarcoidosis; hypersensitivity pneumonitis; multiple sclerosis;
KW allograft rejection; lymphoma; thrombosis; sulphatase;
KW chromosome 20q12-13.2.
XX
XX Synthetic.
FH Key Location/Qualifiers
FT mseq_feature 8160..8701
FT /tag= a
FT /note= "This sequence represents 540 nucleotides not
FT shown in the specification as the result of a printing
FT error"
XX
XX WO200259327-A2.
XX
XX 01-AUG-2002.
XX
XX 26-DEC-2001; 2001WO-US049793.
XX
XX 27-DEC-2000; 2000US-0258577P.
XX
XX 09-FEB-2001; 2001US-0267831P.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Rosen SD, Hemmerich S, Tomita M, Palmeri D;
XX
XX WPI; 2002-636541/68.
XX
XX New sulphatase polypeptides having glucosamine-6-sulphatase activity,
XX useful in screening, discovery and preparation of diagnostic and
XX therapeutic agents for treating cancer, ischemic conditions,
XX inflammation, or thrombosis.
XX
XX Example 5; Page 110-180; 293pp; English.
XX
XX This invention relates to the DNA and protein sequences of a novel
XX polypeptide having glucosamine-6-sulphatase activity. The sulphatases of
XX are useful in screening, discovery and preparation of diagnostic and
XX therapeutic agents for treating cancer, ischemic conditions,
XX inflammation, or thrombosis. The nucleic acids are useful in preparing
XX the sulphatase polypeptides, identifying the expression of genes in a
XX biological specimen, or generating transgenic non-human animals or site-
XX specific gene modification in cell lines. The host cells are useful in
XX replicating and/or expressing the polynucleotides or nucleic acids. The
XX agents are useful in treating the disorders cited above by reducing
XX tumour growth, inflammation, and thrombosis, or increasing angiogenesis,
XX e.g. by treating coronary, carotid, or arterial occlusive disease,
XX peripheral arterial disease, atherosclerosis, myointimal hyperplasia,
XX thromboangiitis obliterans, thrombotic disorders, vasculitis, or
XX preventing ischemic conditions, heart attack (myocardial infarction), or
XX other vascular death. The sulphatases and/or agents are also useful in
XX treating rheumatoid arthritis, asthma, adult respiratory distress
XX syndrome, sarcoidosis, hypersensitivity pneumonitis, multiple sclerosis,
XX allograft rejection, and spread of lymphomas to cutaneous sites. The
XX present sequence represents the human SULF2 genomic DNA sequence of the

CC invention. This gene maps to human chromosome 20q12-13.2
XX
XX SQ Sequence 268685-BP; 72718 A; 63115 C; 64177 G; 67996 T; 0 U; 679 Other;
XX
XX Query Match 5.9%; Score 365.2; DB 6; Length 268685;
XX Best Local Similarity 52.8%; Pred. No. 4,1e-55;
XX Matches 1289; Conservative 0; Mismatches 1043; Indels 108; Gaps 19;
XX
XX 486 GCCTGGCGGCTGGCTGACGCTGTATCCAGCCCTTTGGAGGCCAGATGGCGAGT 545
XX |||||
XX 221964 GCACGGCGGGTGGCTGCTGCTGTATCCCAACTTTGGAGGCCAGATGGCGAGT 222023
XX |||||
XX 546 CATGAGGTGAGAGTTTGGACCAAGCTTGCACATGTTGTAACCCCTCTCTAATA 605
XX |||||
XX 222024 CATGAGGTGAGAGTTTGGACCAAGCTTGCCTTGCATGTTGTAACCCCTCTCTAATA 222083
XX |||||
XX 606 GATACAAAAAATTAGCCAGGTGTGGGGGACACCTGTAAACCCGCTGTGAGAG 665
XX |||||
XX 222084 AATACAAAAAATTAGCTGAGTGTGGGACATCTGTATGCCAGCACTCGGAGGC 222143
XX |||||
XX 666 TGAAGCAGAGAAATGCTTGAACCCAGAGGAGAGG-TGACGCGACAAAGATCAACC 724
XX |||||
XX 222144 TGAGGCAAGAGATGTTGAAACCGGAGGAGAGAGTTGACGTGACCGAGATCTCGCC 222203
XX |||||
XX 725 AATGCACTGATGCTTGTGATGACAGGGCAGACTTCCCACTCAAAAAAAAAAAAAA 784
XX |||||
XX 222204 ACTGCACTGACCTGCGGCAAGAGCGAGCTCGCTCAAAAAAAAAAAAAA 222263
XX |||||
XX 785 AATATGTAAGTTGTGCTATTAACAAATTAATAGAGTGAAGCAAAAGTCTAAAGCTTA 844
XX |||||
XX 222264 AAGCTTGGCCCATATATTTATCTTTCTTATGACATATTCATTAGAAAACATATGAA 222323
XX |||||
XX 845 TGACCAATGTAATCTAGAAATAGCTGGGAAACATATTAAGGAACTTAACCACTGCTG 904
XX |||||
XX 222324 AATCAAGTTTGTGGGAAATCAATAGGCTTCAATGTAAGTAAATCTGCAAA 222383
XX |||||
XX 905 G-AAAGTAAAGTTTGGAAAGAAATGTTGAGACAAAGGTTAAAGAGTAAAAA 963
XX |||||
XX 222384 GCTACTGAGATTTTAAAGGACCTAATATTAAGTGAAGCTGACAGAAATTAACCACTG 222443
XX |||||
XX 964 AATTAATAATACCAAGTTTACTGTGTGAGATAGGGATAGGAGCTAACTAGAAATCAA 1023
XX |||||
XX 222444 CATTAATCTGTTTAAATTCAGAAATTTTAAATGTGACCAAAATCTTATTAGTAGAGG 222503
XX |||||
XX 1024 ATAGCAATGTTTATGATGATGTTAAAGACCTGTGTAAGGTTAAGCACTTACATTATCT 1083
XX |||||
XX 222504 GGAAGCTAGGAAAGGATTAAGCATTCGTGGAAGAGTAATATCTGTGACCTGTG 222563
XX |||||
XX 1084 GCACCATGCGGGACCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1141
XX |||||
XX 222564 CCCAAATAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 222623
XX |||||
XX 1142 AGAATGTAGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
XX |||||
XX 222624 ACAAATACAAAACAGATGCTCAAAAAGTTTAAATGAAACATATGAGGACAGAAAG 222683
XX |||||
XX 1201 AGACAGTGAAGGAGAGAGAGTTTAAAGATTTATCAAGACAAAGTTAATG-TGCTGGCCAG 1259
XX |||||
XX 222684 ATAAATATATCCCAAGAGGCAAAAGTAAGTAGATCTGATGATGATGATGATGATGATG 222743
XX |||||
XX 1260 TGGCATCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1319
XX |||||
XX 222744 ACAAAGCATTTATCTCTCTCTGCAATTTTTCCTCAATGCTCTCTCTCTCTCTCT 222803
XX |||||
XX 1320 TTATGAAGAGTGAAGGCTTCAAGACCTTCAAGGTGTGATTAATGAGAAATGATT 1379
XX |||||
XX 222804 CATGCTGTATTTCTTAATCTTTTAAATGAACATGATCTTTTAAATGATGATGATGATG 222863
XX |||||
XX 1380 GGAGAAAGATTAATCTGTGAGAGATGAGATTTTCAAGGCTTGAAGTCACTCAATCCAG 1439
XX |||||
XX 222864 TAATTTTCA-----AGATGTACTGTGTAATGCTTAAGATATTTT 222907
XX |||||
XX 1440 AACTGTGCTAAATGCTTCAAGACATGATCTCTGCTCAAGGACCTTAAGCCAAAAA 1499
XX |||||

Db 222908 ATTGATGATCTTCGATGCAATATGTAACATGGCTTTCAAGTAGAATAATTAACACTGC 222967
Qy 1500 CAAGAGATGAAGAAATACACCAATACTATATAGCACTTGTGTGAATATCAAGAAA 1559
Db 222968 CAATATCAATAGAGGTGACAGAAAGACT---CACTCTAAAGATATAATATATATTA 223023
Qy 1560 GAAATACGATCTGATCTGTATGTGTCAACGAGATCAAGATATCTTCTAGTTCAAGA 1619
Db 223024 CAATGCTGTCTACTATGTATTTATCTGAGGGAGTTAAATACAGAGAAAATGTTCAATTTAA 223083
Qy 1620 AGTTTCAGATCGGCGGCGGCGGTGCTCAGCCTGTATATCCAGCACTTTGGAGGCGG 1679
Db 223084 AGTAACATTA-CAGCGGGCATGTGTGCTCAATGCTGTATATCCAAACTTTGGAGGCGG 223142
Qy 1680 AGCGGGGTGATCAC--AAGTCAGAGATCAAGACCATCTGTGTAAACAGGTGAAACC 1737
Db 223143 AGGCAAGCGGATTAATCTGAGGTCAGGACTTTGAGACCAAGCCTGACCAATTTGTGAACC 223202
Qy 1738 CCGTCTTACAAAATAATTAATAAATTAAGCAGGCGGTGTGGGGGCGCTGTAGTCCCA 1797
Db 223203 CTGTCTTACTTAAATAATGTAAATAATTAAGCAGGCGGTGTGGGGCGCTGTAAATTTCA 223262
Qy 1798 GCTTACGAGAGGCTGAGGAGAGAGATGAGCGTGAACCCGGAGGTAGATTGGC-GTGA 1856
Db 223263 GCTACTGGGAGGCTGAGGAGAGATCACTTGAACCCAGAGGAGAGAGGTTGCAAGTGA 223322
Qy 1857 GCGGAGATCGCGCCACTGCGCTCCAGCTTGGG--CGACAGAGTAGACTGCGTCTCAAAA 1914
Db 223323 GCTGAGATCATGCGCATGTCATCTCAGCGCTGGGCGACGACAGGCAAAACTCCGCTCAAAA 223382
Qy 1915 AAAAAAAAAAAAAAAAAAGTTTCAGATCTTAAACAATGCACTTTCACAGCTTAGAA 1974
Db 223383 AAAAAAAAAAGTAAACA-----TCCATTCAGCTTAAC 223413
Qy 1975 TAGAGAGCATGTTACAGGAGAGAAATGTTTTCAGCAAAAGTATACAGTAGGAAATA 2034
Db 223414 AAGTTATACTGTGAATGTGCAAAATATCTTTGCTACTTTAATTTGTGAATAATTAATC 223473
Qy 2035 GAGGATATGTTCAAGAGAGAGACCCAGAGTCAATGTTTGTAGGTTAGAGAAAACA- 2093
Db 223474 TCAAAATCTGAAAGATAGGTTCTGATGCAATTAATGATTAATGACTGCAAAATAG 223533
Qy 2094 CAGTGTTCGCAATCTCAGGTTCCATTAAGGTTATGAATCAATATGCTGTAGCA 2153
Db 223534 CAATATTTCTTTAAAAATATCTGCAATTATGTATATATATATATATTAACAAACACAC 223593
Qy 2154 ACCTGCAATTTAAAAATGAATTAATGATGAGAGAGATGAATAATTTAGCATGCA 2213
Db 223594 ACATACATATGTATATGTAGTAGTATTTTATGTATACCTTCCATTTT---TATA 223649
Qy 2214 TTACATTTTGAAGAGCAATATATTTTTCGCACTTTGCTCAATTTGAATCTGTAAT 2273
Db 223650 TCACTTGTAAAGTCCACATTTGAAATTAAG-----TGTTCACAAAGCA 223695
Qy 2274 TATATTTTATGATGATGTGAATACAGATACATATATATTTCTTACTGTAGCATGCA 2333
Db 223696 AAGATATACGGCTTAAGTCAAGGAGACTGAGCTTAATTTTCTTATATAGCATGATCTTG 223755
Qy 2334 GTCAAAAAATCTTTAAAGCATGGCTGTCTTAATCTTATTTTTCAGAGAGAGAAATC 2393
Db 223756 GACAATTAACCAATCTGTGCAACCAAGCTTCTGTTCTTATATCTGAGATGGGGCA--- 223812
Qy 2394 CAGATCTGAGAGCAAAACATTTTGCCTGAGGTTATGAACAGCTTATGCAATGCTTA 2453
Db 223813 -----TATATGACAAACCCCTTGCAATCCCATTTACCTTAGTGACATTAAGTA 223864
Qy 2454 AAGATGATCTTATGTTAAATTTCTTTCCAGTGTGCAATGCACTTCTAGTCTGTT 2513
Db 223865 ACCCTGAAGTTAAATTAATATATTTGCAAGCTGTTTATTTCAATATGATCATCTTAT 223924
Qy 2514 GGCCTGAATATACGAATATATTTAGTGAACAGATACACAGTCTGGGGAATATATTTGG 2573

Db 223925 CAACCAAGATGCTTTAATTTT-----TATTTCTATGATTAATAAATGG 223969
Qy 2574 GTAGTGCGTGAAGAGCCTCATTTTCTPAAGAAATGTGAACCTTAGCAGGGTATGTGGCT 2633
Db 223970 CAGACCAAGATTGAAACTCAGGTTGACGTCAAGAAATTACTCCCGGCCAAGCACAGTGCT 224029
Qy 2634 CACACCTATTAATTTCCAGCACTTTGGAGGCGCAAGTCAAGAAATGCTGTAACCCAAAG 2693
Db 224030 CACACTGTAAATCCAGCAATTTTGGAGGCTGAGGCGAGCACTCACTTAGGCTCAAGAG 224089
Qy 2694 TTCAAGACTAGCATGAGGCAACAATAGCAAGACCTCATCTCTACAAAAAATTTAAATTCAG 2753
Db 224090 TTGGAACCAAGCTGAGCAACAATGTGGAACCCCATCTCTATTAATAATCAAAATTTAG 224149
Qy 2754 CTGAGCATGTTGGCATACGCTGTAGTCCACCTTACTTGGGAAAGCTTA-GTGGGTGTGATC 2812
Db 224150 CCGGCGATGTTGGCAGCTGCTGTAGTCCAACTGTTTGGAGGCCATGCGAGAAATTC 224209
Qy 2813 GCTTG-ACACAGAGTTTGAAGCTAAGTGAGCCATGATCACACAACTGCACTCCAGCTT 2871
Db 224210 ACTTGAAACCAAGAGTGAAGGTTGCAATGAGCCGAGATCGCACTTGCACAGCTT 224269
Qy 2872 GAGTGACAGAGGAAGACCTGCTCCCTTAAAAAAGAAAGAA 2911
Db 224270 GGGCAACAGAGCAAGACTCCGCTCTCAACAAAAAAGAAAAA 224309

RESULT 11
AAH72636
ID AAH72636 standard; cDNA; 2724 BP.
XX
XX AAH72636;
XX
DT 19-SEP-2001 (first entry)
XX
XX Human cervical cancer marker nucleic acid 3910.
XX
XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
XX Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000MO-US033312.
XX
XX 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0189315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
XX WPI; 2001-375006/39.
DR
PT New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
XX
PS Claim 1; Page 774; 1051p; English.
XX
XX The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytosarstatic activity. The nucleic acids and encoded
CC polypeptides are useful; to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy

XX Sequence 2724 BP; 906 A; 552 C; 605 G; 661 T; 0 U; 0 Other;
SQ
Query Match 5.5%; Score 338; DB 4; Length 2724;
Best Local Similarity 88.0%; Pred. No. 1.8e-50;
Matches 368; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 16 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGTGAGCTTAATATCCAGAG 75
DB 332 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGTGAGCTTAATATCCAGAG 391
QY 76 AACTCCCTTTGTAAGATTGTGACAAAATTAATGATGAGTAAATAGTCTTAAG 135
DB 392 CACCCCTTTGTAAGATTGTGACAAAATTAATGATGAGTAAATAGTCTTAAG 451
QY 136 GATGTGTAACCCAGAGCATATCAGCCCTAGCAAAATGCGAATTCATATATCATCA 195
DB 452 GAGTGTGACCCAGAGCATATCAGTCTAGCAAAATGCGAATTCATATATCATCA 511
QY 196 AAGTTATCTTCAAGAGCTTCAGCCCTTAATGATGTCTTAAGAAAATGTGAAAGCCCTC 255
DB 512 AAGTTGTCTTCAAGAGCTTCAGCCCTTAATGATGTCTTAAGAAAATGTGAAAGCCCTC 571
QY 256 AGCCATCTGAGAGACAGTGTACAGCAATGATCAAAAAGAAAACACAGGCCCTTCCC 315
DB 572 AGCCGCCGAGAGACAGTGTACAGCAATTAATCAAAAAGAAAACACAGGCCCTTCCC 631
QY 316 CTTCCCCCACTGATGATGAGAGAGTCTTCATTTCCATAGTAAATTTTCTAGATAC 375
DB 632 CTTCCCCCAATTCATGATGATGAGAGAGTCTTCATTTCCATAGTAAATTTTCTAGATAC 691
QY 376 AGCTGTAGAGCTCAAAAGTACTGAGAAAAGAGCTCCCATTCAGAGAAATTTATCTTA 433
DB 692 GTCTGTAGAGCTCAAAAGTACTGAGAAAAGAGCTCCCATTCAGAGAAATTTATCTCA 749

RESULT 12
AAH72784
ID AAH72784 standard; cDNA; 2725 BP.
XX
AC AAH72784;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 4058.
XX
KW Human cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US033312.
XX
PR 08-DEC-1999; 99US-0169681P.
XX
PR 21-DEC-1999; 99US-0171350P.
XX
PR 14-MAR-2000; 2000US-0189315P.
XX
PR 12-MAY-2000; 2000US-0203791P.
XX
PR 09-JUN-2000; 2000US-0210600P.
XX
PR 21-JUL-2000; 2000US-0220114P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deede J, Berger A, Zhao X;
XX
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer and
XX
XX PT for assessing and detecting compounds for treating the cancer.
XX
XX PS Claim 1; Page 850; 1051p; English.

XX The invention relates to novel genes (AAH6727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
SQ Sequence 2725 BP; 906 A; 552 C; 605 G; 662 T; 0 U; 0 Other;
Query Match 5.5%; Score 338; DB 4; Length 2725;
Best Local Similarity 88.0%; Pred. No. 1.8e-50;
Matches 368; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 16 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGTGAGCTTAATATCCAGAG 75
DB 332 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGTGAGCTTAATATCCAGAG 391
QY 76 AACTCCCTTTGTAAGATTGTGACAAAATTAATGATGAGTAAATAGTCTTAAG 135
DB 392 CACCCCTTTGTAAGATTGTGACAAAATTAATGATGAGTAAATAGTCTTAAG 451
QY 136 GATGTGTAACCCAGAGCATATCAGCCCTAGCAAAATGCGAATTCATATATCATCA 195
DB 452 GAGTGTGACCCAGAGCATATCAGTCTAGCAAAATGCGAATTCATATATCATCA 511
QY 196 AAGTTATCTTCAAGAGCTTCAGCCCTTAATGATGTCTTAAGAAAATGTGAAAGCCCTC 255
DB 512 AAGTTGTCTTCAAGAGCTTCAGCCCTTAATGATGTCTTAAGAAAATGTGAAAGCCCTC 571
QY 256 AGCCATCTGAGAGACAGTGTACAGCAATGATCAAAAAGAAAACACAGGCCCTTCCC 315
DB 572 AGCCGCCGAGAGACAGTGTACAGCAATTAATCAAAAAGAAAACACAGGCCCTTCCC 631
QY 316 CTTCCCCCACTGATGATGAGAGAGTCTTCATTTCCATAGTAAATTTTCTAGATAC 375
DB 632 CTTCCCCCAATTCATGATGATGAGAGAGTCTTCATTTCCATAGTAAATTTTCTAGATAC 691
QY 376 AGCTGTAGAGCTCAAAAGTACTGAGAAAAGAGCTCCCATTCAGAGAAATTTATCTTA 433
DB 692 GTCTGTAGAGCTCAAAAGTACTGAGAAAAGAGCTCCCATTCAGAGAAATTTATCTCA 749

RESULT 13
ABV30259
ID ABV30259 standard; cDNA; 2725 BP.
XX
AC ABV30259;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 30250.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
XX
PR 16-MAR-2000; 2000US-0189862P.
XX
PR 25-MAY-2000; 2000US-0207454P.
XX
PR 09-JUN-2000; 2000US-0211314P.
XX
PR 18-JUL-2000; 2000US-0219007P.
XX
PR 13-DEC-2000; 2000US-0255281P.
XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

PT Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 6566; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 2725 BP; 905 A; 552 C; 606 G; 662 T; 0 U; 0 Other;

Query Match 5.5%; Score 338; DB 5; Length 2725;

Best Local Similarity 88.0%; Pred. No. 1.8e-50; Matches 368; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 16 CAATTATGAAAGCCAAATATACAGCTTAATGAAATGAGCTTAATACCCGAAG 75

Db 332 CAATTATGAAAGCCAAATATACAGCTTAATGAAATGAGCTTAATACCCGAAG 391

QY 76 AACTCCCTTTGTAAGATTGTAACAAAATTAATAGATTAAGTTCTAATG 135

Db 392 CACCCCTTTGTAAGATTGTAACAAAATTAATAGATTAAGTTCTAATG 451

QY 136 GAATGTGAACCCAGACCATATACAGCTTACGAAATGCGAATTCATATATCA 195

Db 452 GAGTGTGAGACCCAGACCATATACAGCTTACGAAATGCGAATTCATATATCA 511

QY 196 AAGTTATCCTTGAAGAGCTTACAGGCTTATGATGCTTAAGAAATGGAAGCCCTC 255

Db 512 AAGTTATCCTTGAAGAGCTTACAGGCTTATGATGCTTAAGAAATGGAAGCCCTC 571

QY 256 AGCCATCTGAAGAGAGTGTTCAGCAATGATCAAAAAGAACCCAGGCTTCCC 315

Db 572 AGCCGCGGAGAGAGTGTTCAGCAATGATCAAAAAGAACCCAGGCTTCCC 631

QY 316 CTTCCCTTGAAGAGTGTTCAGCAATGATCAAAAAGAACCCAGGCTTCCC 375

Db 632 CTTCCCTTGAAGAGTGTTCAGCAATGATCAAAAAGAACCCAGGCTTCCC 691

QY 376 AGCTTGAAGAGTCAAGTACTGAGAAAGAGCTCCATTCAAGGAATTTATCTTA 433

Db 692 GCTTTGTAAGAGTCAAGTACTGAGAAAGAGCTCCATTCAAGGAATTTATCTTA 749

RESULT 14

ABV27562 ABV27562 standard; cDNA; 2725 BP.

XX ABV27562;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 27553.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX

OS Homo sapiens.

PN WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183119P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0235281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 5630-5631; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the efficacy of a test compound to inhibit prostate cancer in a patient; (c) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (d) assessing the efficacy of a composition for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 2725 BP; 905 A; 552 C; 606 G; 662 T; 0 U; 0 Other;

Query Match 5.5%; Score 338; DB 5; Length 2725;

Best Local Similarity 88.0%; Pred. No. 1.8e-50; Matches 368; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 16 CAATTATGAAAGCCAAATATACAGCTTAAATGAAATGAGCTTAATACCCGAAG 75

Db 332 CAATTATGAAAGCCAAATATACAGCTTAAATGAAATGAGCTTAATACCCGAAG 391

QY 76 AACTCCCTTTGTAAGATTGTAACAAAATTAATAGATTAAGTTCTAATG 135

Db 392 CACCCCTTTGTAAGATTGTAACAAAATTAATAGATTAAGTTCTAATG 451

QY 136 GAATGTGAACCCAGACCATATACAGGCTTACGAAATGCGAATTCATATATCA 195

Db 452 GAGTGTGAGACCCAGACCATATACAGGCTTACGAAATGCGAATTCATATATCA 511

QY 196 AAGTTATCCTTGAAGAGCTTACAGGCTTATGATGCTTAAGAAATGGAAGCCCTC 255

Db 512 AAGTTATCCTTGAAGAGCTTACAGGCTTATGATGCTTAAGAAATGGAAGCCCTC 571

QY 256 AGCCATCTGAAGAGAGTGTTCAGCAATGATCAAAAAGAACCCAGGCTTCCC 315

Db 572 AGCCGCGGAGAGAGTGTTCAGCAATGATCAAAAAGAACCCAGGCTTCCC 631

QY 316 CTTCCCTTGAAGAGTGTTCAGCAATGATCAAAAAGAACCCAGGCTTCCC 375

Db 632 CTTCCCTTGAAGAGTGTTCAGCAATGATCAAAAAGAACCCAGGCTTCCC 691

QY 376 AGCTTGAAGAGTCAAGTACTGAGAAAGAGCTCCATTCAAGGAATTTATCTTA 433

Db 692 GCTTTGTAAGAGTCAAGTACTGAGAAAGAGCTCCATTCAAGGAATTTATCTTA 749

RESULT 15
 ID ABV21740 standard; cDNA; 2725 BP.
 XX ABV21740
 AC ABV21740;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 21731.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 XX
 PN MO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX
 PR 17-FEB-2000; 2000US-018319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI: 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 3668-3669; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (1) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (1) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 2725 BP; 905 A; 552 C; 606 G; 662 T; 0 U; 0 Other;
 XX
 Query Match 5.5%; Score 338; DB 5; Length 2725;
 Best Local Similarity 88.0%; Pred. No. 1,8e-50;
 Matches 368; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 QY 16 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGTGACCTAAATACCCGAAG 75
 DB |||||
 QY 332 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGTGACCTAAATACCCGAAG 391
 DB |||||
 QY 76 AACTCCCTTTTGAAGATTGTGAACAAATTAATTAATGAGTTAATAGTTCTTAATG 135
 DB |||||
 QY 392 CACCCCTTTTGAAGATTGTGAACAAATTAATTAATGAGTTAATAGTTCTTAATG 451
 DB |||||
 QY 136 GATGTGTAACCAAGACCATATGACGCTAGCAAAATGCGAATTCATATATCATCA 195
 DB |||||
 QY 452 GAGTGTGTAACCAAGACCATATGACGCTAGCAAAATGCGAATTCATATATCATCA 511
 DB |||||

QY 196 AAGTTATCTTCAAGAGCTTCAGCCCTTAATATGTCTTAAAGAAATGTGAAACGCCCTC 255
 DB |||||
 QY 512 AAGTTATCTTCAAGAGCTTCAGCCCTTAATATGTCTTAAAGAAATGTGAAACGCCCTC 571
 DB |||||
 QY 256 AGCCATCTGAAGACAGTGTTCAGCAATTTGATCAAAAAAGAAAAACACAGGCCCTTCCC 315
 DB |||||
 QY 572 AGCCGCCGGAAGACAGTGTTCAGCAATTTGATCAAAAAAGAAAAACACAGGCCCTTCCC 631
 DB |||||
 QY 316 CTTCCCTCACTTGAATGATGACAGCTTTCATTTTCCATAGTATTAATTTTCTAGATAC 375
 DB |||||
 QY 632 CTTCCCTCACTTGAATGATGACAGCTTTCATTTTCCATAGTATTAATTTTCTAGATAC 691
 DB |||||
 QY 376 AGCTGTGATGCTCAAAAGTACTGGAAGAAAGCTCCCATTCAGAAAGAAATTTATCTTA 433
 DB |||||
 QY 692 GTCTTGTAGACCTCAAAAGTACTGGAAGAAAGCTCCCATTCAGAAAGAAATTTATCTTA 749
 DB |||||

Search completed: December 26, 2004, 18:49:00
 Job time : 1806.57 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2004, 10:56:59 ; Search time 22 Seconds
(without alignments)
1902.468 Million cell updates/sec

Title: US-10-786-065-5

Perfect score: 2320

Sequence: 1 MQGELCAKTVPGSCCYHCS.....CDLLASYGKHHPAPSKCW 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	916	39.5	496	2 S23385	protein kinase (EC
2	912	39.3	496	2 S30435	protein kinase PCT
3	908	39.1	496	2 JCS110	cyclin-dependent k
4	907.5	39.1	461	2 JCS111	cyclin-dependent k
5	900	38.8	451	2 S30436	protein kinase PCT
6	897	38.7	523	2 S23384	protein kinase (EC
7	876.5	37.8	380	2 S32831	serine/threonine-s
8	819	35.3	378	2 B87722	protein ZC123.4 [I
9	800.5	34.5	577	2 T15445	hypothetical prote
10	745.5	32.1	292	2 S40021	protein kinase (EC
11	745	32.1	302	1 OKB85	protein kinase PHO
12	743	32.0	292	1 A45091	protein kinase (EC
13	743	32.0	292	1 A45092	protein kinase (EC
14	743	32.0	293	2 JE0374	cyclin-dependent c
15	739	31.9	288	2 T41101	protein kinase (EC
16	737	31.8	292	1 A46365	protein kinase (EC
17	734	31.6	305	2 S23386	protein kinase (EC
18	731.5	31.5	291	2 S23386	protein kinase (EC
19	714	30.8	298	1 A44878	protein kinase (EC
20	713	30.7	288	1 A42566	protein kinase (EC
21	701	30.2	294	2 S51008	protein kinase (EC
22	701	30.2	296	2 S24386	protein kinase (EC
23	695.5	30.0	294	2 S23095	protein kinase (EC
24	694.5	29.9	294	1 S42049	protein kinase (EC
25	691.5	29.8	294	1 B40444	protein kinase (EC
26	691.5	29.8	294	2 S22440	protein kinase (EC
27	691.5	29.8	302	2 T17115	protein kinase cdc
28	690	29.7	308	1 S53538	protein kinase (EC
29	689.5	29.7	294	1 A40444	protein kinase (EC

30	689.5	29.7	294	2 T49271	CELL DIVISION CONT
31	688	29.7	292	2 S22441	protein kinase (EC
32	686	29.6	292	2 T25374	hypothetical prote
33	683.5	29.5	294	1 JQ2243	protein kinase (EC
34	681.5	29.4	291	2 A39107	protein kinase (EC
35	679.5	29.3	294	1 S57928	protein kinase (EC
36	679	29.3	298	2 A41227	protein kinase (EC
37	671	28.9	301	1 A48041	protein kinase (EC
38	663	28.6	293	2 T02922	protein kinase (EC
39	662.5	28.6	317	1 JC4827	protein kinase (EC
40	662	28.5	297	2 A37871	protein kinase (EC
41	662	28.5	301	1 S42101	protein kinase (EC
42	658.5	28.4	294	1 S31332	protein kinase (EC
43	652	28.1	346	1 T78840	protein kinase (EC
44	650	28.0	302	1 I50474	protein kinase (EC
45	646	27.8	346	1 I48157	protein kinase (EC

ALIGNMENTS

RESULT 1

S23385 protein kinase (EC 2.7.1.37) cdc2-related PCTAIRE-1 - human

C/Species: Homo sapiens (man)

C/Date: 18-Jun-1993 #sequence_revision 18-Jun-1993 #ext_change 16-Aug-2004

C/Accession: S23385; S22747

R/Meyerson, M.; Ender, G.H.; Wu, C.L.; Su, L.K.; Gorka, C.; Nelson, C.; Harlow, E.; Tsai

EMBO J. 11, 2909-2917, 1992

A/Title: A family of human cdc2-related protein kinases.

A/Reference number: S23382; PMID:92347325; PMID:1639063

A/Accession: S23385

A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-496 <MEY>

A/Cross-references: UNIPROT:000536; EMBL:X66363; NID:G36618; PIDN:CAA47006.1; PID:G36619

C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F/163-415/Domain: protein kinase homology <KIN>

F/171-179/Region: protein kinase ATP-binding motif

Query Match	39.5%	Score 916;	DB 2;	Length 496;
Best Local Similarity	44.7%	Pred. No. 5.8e-35;		
Matches	202;	Conservative	67;	Mismatches 121; Indels 62; Gaps 9;
QY	23	GENHSC---RSQPTTEAFAKL-TDLKASCSMTSFHRCGLQAAQFKSKRPSNSD	78	
DB	54	GELRSARGLSSAPETVHEDLMGSDGSDQASATSSDE--VQSPVRVAMRHHPRKIS-	110	
QY	79	CFQEDDLRGQFQWRKSLP-----FGAAS	101	
DB	111	---TEDINK---RLSLPADIRLPEGYLEKLTLSNPIDPKPLRRRLRVLSLSEIGFKUE	163	
QY	102	SVLNLEKLGEGSVATYVKISIRINGQVLAKYISMAEAGVFTAIRASLKLGLGHANV	161	
DB	164	TYIKDKRGEGYATYVYKSKSLTDNLVAKKIRLHEHSGACCTAIRVSLDLGLHANI	223	
QY	162	VLLHDIHTKETLTVEFEYMTDLAQYMSQHGGGLPHNVRLFMFQLNLGLAYIHHQVYL	221	
DB	224	VTLHDIHTKESLTIVFEYLDKQYLDGCGNIINMHNVKFLPQLAGLAVCHROKVL	283	
QY	222	HRDLKQNLILHGLGLKADPGLARAKSIPTQYTSSEVYTLWYRPDALLGATYSSGL	281	
DB	284	HRDLKQNLILHGLGLKADPGLARAKSIPTQYTSSEVYTLWYRPDALLGATYSSGL	343	
QY	282	DLMGAGCPIEMFGQGLPFGVSNLLEOLEKMEVLTGVTEDTPGWGSLPVYNPEWPL	341	
DB	344	DMWGAGCPIEMFGQGLPFGVSNLLEOLEKMEVLTGVTEDTPGWGSLPVYNPEWPL	402	
QY	342	PPPRSLHVVNRLGRVP---EAEDLASQMLKGFPRDRVSAQEAIVHDYFSLAPSQLYOL	397	
DB	403	YAAEAL-----LSHAPRLDSGADLLTKLTLFEGGNRISADAMKHPFLSGEIRIHL	456	

QY 398 PDESLFTVSGVRLKPEKMDLASYQKGNHPA 429
 Db 457 PDTTISIFALKEIQLQKASLSRSSMPSDGRPA 488

RESULT 2

S30435
 protein kinase PCTAIR-1 (EC 2.7.1.-) - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Aug-2004
 C/Accession: S30435
 R/Okuda, T.; Cleveland, J.L.; Downing, J.R.
 Oncogene 7, 2249-2258, 1992
 A>Title: PCTAIR-1 and PCTAIR-3, two members of a novel cdc2/CDC28-related protein kinase
 A/Reference number: S30435; MUID:93064701; PMID:11437147
 A/Accession: S30435
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-496 <OKU>
 A/Cross-references: UNIPROT:Q04735; EMBL:X69025; NID:953610; PIDD:CAA48787.1; PID:953611
 C/Genetic: 68/1
 A/Introns: 68/1
 C/Superfamily: protein kinase homology
 C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 F/163-415/Domain: protein kinase homology <KIN>
 F/171-179/Region: protein kinase ATP-binding motif

Query Match 39.3%; Score 912; DB 2; Length 496;
 Best Local Similarity 45.8%; Pred. No. 8.8e-55;
 Matches 200; Conservative 64; Mismatches 111; Indels 62; Gaps 9;

QY 23 GEASRC---RSQPTTEAFKL-TDLKASCSMTSFHPRGLQAAQAQKFKSKRRSNSD 78
 Db 54 GELRSVRGPISSAPETVHEDMKSGSDQASATSDS--VQSPVRVRRMRNHPRKIS- 110
 QY 79 CFEEDLRQGFQWRKSLP-----FGAAS 101
 Db 111 ---TEDINK---RLSLPADIRLPEGLYEKLTINSPIFDKPLSRRLRVSLSIGGCKE 163
 QY 102 SYINLEKLGSGYATYTKGIRINGQVALKVISMAEBCVPPTAIRASLLKGLKHANI 161
 Db 164 TYTKDKLGEGYATYTKGSKLTDNLVVLKEIRLHEEGAPCTAIRVSLKDLKHANI 223
 QY 162 VLLHDIITTKETLTFFVEYVHTDIAQYMSQHPCGLPHNVRLEMPOLAGLYIHQVHL 221
 Db 224 VTLHDIITHEKSLTVFELDKQYLDDCGNVIMHNVKLPFLGLGLAACHQKVL 283
 QY 222 HRDLKQNLILSHLGLKADFGIARAKSIPSQYSSVVTLWYRPDALLGATEYSEL 281
 Db 284 HRDLKQNLILINERGLKADFGIARAKSIPTKYSNEVVTLWYRPDILLSTQSTOI 343
 QY 282 DIMGAGCITIEHQGPPLFPQVSNILKQLEKIEWLGVPTEDTDPGVSKLPVNPWFPL 341
 Db 344 DMVGVCITIEYKATGKPLFPFG-STVEEQHLFIIRLIGTPTLEETPGILSNEERTYNYEK 402
 QY 342 PTPRSIHVVWNRIGRPV---EAEDLASQMLKGFPRDRVSAOBLVHDYFSLPSQLYL 397
 Db 403 YAAEAL-----LSHAPRLDSOGADLLTTLQLPREGNKRISAEDAKHPPFLSLGERIKHL 456
 QY 398 PDESLFTVSGVRLKPE 414
 Db 457 PDTTISIFALKEVQLQKE 473

RESULT 3

UC5110
 cyclin-dependent kinase-related protein 1a - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Aug-2004
 C/Accession: UC5110
 R/Gao, C.Y.; Chauthaiwale, V.M.; Rampalli, A.M.; Zelenka, P.S.
 Gene 176, 243-247, 1996
 A>Title: Expression of alternatively spliced PCTAIR-1 mRNA in PC12 cells and neonatal

A/Reference number: UC5110; MUID:97075937; PMID:8918260
 A/Contents: PC12 cell
 A/Accession: UC5110
 A/Molecule type: mRNA
 A/Residues: 1-496 <GAO>
 A/Cross-references: UNIPROT:Q63686; GB:U36444; NID:91040966; PIDD:AAC52912.1; PID:9104096
 C/Comment: This protein plays a role in differentiation and in apoptosis of differentiated
 C/Superfamily: protein kinase homology
 C/Keywords: ATP
 F/163-415/Domain: protein kinase homology <KIN>
 F/171-179/Region: protein kinase ATP-binding motif

Query Match 39.1%; Score 908; DB 2; Length 496;
 Best Local Similarity 46.6%; Pred. No. 1.3e-34;
 Matches 200; Conservative 65; Mismatches 118; Indels 46; Gaps 10;

QY 23 GEASRC---RSQPTTEAFKL-TDLKASCSMTSFHPRGLQAAQAQKFKSKRRSNSD 77
 Db 54 GELRSVRGPISSAPETVHEDMKSGSDQASATSDS--VQSPVRVRRMRNHPRKIST 111
 QY 78 ----DQFQ--EEDLR--QGFQWRKSLP-----FGAASVYNLEKL 109
 Db 112 EDINKLSLPADIRLPEGLYEKLTINSPIGDKPLSRRLRVSLSIGGCKETTYKDL 171
 QY 110 GEGSYATYTKGIRINGQVALKVISMAEBCVPPTAIRASLLKGLKHANIVLLHDIH 169
 Db 172 GEGYATYTKGSKLTDNLVVLKEIRLHEEGAPCTAIRVSLKDLKHANIVTLHDIH 231
 QY 170 TKEITLTPVEYVHTDIAQYMSQHPCGLPHNVRLEMPOLAGLYIHQVHLHPLKPN 229
 Db 232 TEKSLTVFELDKQYLDDCGNVIMHNVKLPFLGLGLAACHQKVLHPLKPN 291
 QY 230 ILISHGEIKLADFGIARAKSIPSQYSSVVTLWYRPDALLGATEYSELDIMGAGCI 289
 Db 292 ILINERGLKADFGIARAKSIPTKYSNEVVTLWYRPDILLSTQSTQIDMVGVC 351
 QY 290 PTEMQGPPLFPQVSNILKQLEKIEWLGVPTEDTDPGVSKLPVNPWFPLPTPRSLHV 349
 Db 352 FYEMATGKPLFPFG-STVEEQHLFIIRLIGTPTLEETPGILSNEERTYNYKYAEAL-- 408
 QY 350 VWNRLGRVPEAE---DLASQMLKGFPRDRVSAOBLVHDYFSLPSQLYLPDESLFT 405
 Db 409 ----LRHAPRLDSOGADLLTTLQLPREGNKRISAEDAKHPPFLSLGERIKHLPTTSTIFA 464
 QY 406 VSGVRLKPE 414
 Db 465 LKEVQLQKE 473

RESULT 4

UC5111
 cyclin-dependent kinase-related protein 1b - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 16-Aug-2004
 C/Accession: UC5111
 R/Gao, C.Y.; Chauthaiwale, V.M.; Rampalli, A.M.; Zelenka, P.S.
 Gene 176, 243-247, 1996
 A>Title: Expression of alternatively spliced PCTAIR-1 mRNA in PC12 cells and neonatal re
 A/Reference number: UC5110; MUID:97075937; PMID:8918260
 A/Contents: PC12 cell
 A/Accession: UC5111
 A/Molecule type: mRNA
 A/Residues: 1-461 <GAO>
 A/Cross-references: UNIPROT:Q63686; GB:U36444; NID:91040966; PIDD:AAC52913.1; PID:9104096
 C/Comment: This protein plays a role in differentiation and in apoptosis of differentiated
 C/Superfamily: protein kinase homology
 C/Keywords: ATP
 F/128-380/Domain: protein kinase homology <KIN>
 F/136-144/Region: protein kinase ATP-binding motif

Query Match 39.1%; Score 907.5; DB 2; Length 461;
 Best Local Similarity 55.6%; Pred. No. 1.3e-34;
 Matches 179; Conservative 52; Mismatches 80; Indels 11; Gaps 3;

```

QY 97 FGAASYLNLKLEGGSYATVYKGISRINGQVALKVISMNAEAGVPFPAIRBASLLKGL 156
DB 124 FGLLETVYKLDKLEGGSYATVYKGSKLTIDNLVALKEIRLEHEEGAPCTAIRREVSLIKDL 183
QY 157 KANITVTLHDIIHTKKTLLFVFEVYMHDTLAQVMSQHPGLHPHNVRLLPFWOLLRGLAYTH 216
DB 184 KANITVTLHDIIHTKKTLLFVFEVYLDKDLKQYDDCGNVTNMKNVTLFLFOLLRGLAYCH 243
QY 217 HOHVLHRDLKPOLNLTISHGELKLDADFGIARAKSIPSQTSSESVTLWYRPDALLGATE 276
DB 244 RQKVHLRDLKPOLNLTINRGELKLDADFGIARAKSIPTKTYSNVVTLWYRPDILLGSTD 303
QY 277 YSESLDIWAGCIFIEMFGQPLPFGVSNILEOLEKIWEVLGVPTEDTWPVGSKLPNYNP 336
DB 304 YSQIDIMWGVCIFYEMATGRPLFPFG-STVEEDLHIFRLLGTPTEDTWPGLISNEEFPRT 362
QY 337 EWPPLPPTPRSLHVMNRIGRVPFAE---DLASQMLKGFPRDRVSQGLVHDFYSLALS 392
DB 363 YNPYKTRABAL-----LHAPRLBECGDADLTLYLLQFEGRNRIISADNAKHPFLSLGB 416
QY 393 QLYQLDEESLFTVSGVRLKPE 414
DB 417 RIHKLPDTTISIFALKVQLAKE 438

RESULT 5
S30436
protein kinase PCTAIR-3 (EC 2.7.1.1.-) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Aug-2004
C:Accession: S30436
R:Okuda, T.; Cleveland, J.L.; Downing, J.R.
OncoGene 7, 2249-2258, 1992
A>Title: PCTAIR-1 and PCTAIR-3, two members of a novel cdc2/cdc28-related protein kinase
A:Reference number: S30435; MUID:93064701; PMID:1437147
A:Accession: S30436
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-451 <OKU>
C:Cross-references: UNIPROT:Q04899; EMBL:X69026; NID:953612; PIDN:CAA48788.1; PID:953613
C:Superfamily: protein kinase homology
C:Keyword: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:119-371/Domain: protein kinase homology <KIN>
F:127-135/Region: protein kinase ATP-binding motif

Query Match 38.8%; Score 900; DB 2; Length 451;
Best Local Similarity 49.9%; Pred. No. 2.8e-34;
Matches 190; Conservative 54; Mismatches 87; Indels 50; Gaps 8;

QY 74 RNSNDPFOBEDLRQGFQWRKSLP----- 96
DB 59 RQNRFRSMEDLNK-----RLSLPMDIRLPQEFLOKLQLENPGLPKPLTMSRRAASLSDIG 114
QY 97 FGAASYLNLKLEGGSYATVYKGISRINGQVALKVISMNAEAGVPFPAIRBASLLKGL 156
DB 115 FGLLETVYKLDKLEGGSYATVYKGSKLTIDNLVALKEIRLEHEEGAPCTAIRREVSLIKDL 174
QY 157 KANITVTLHDIIHTKKTLLFVFEVYMHDTLAQVMSQHPGL-HPHNVRLLPFWOLLRGLAYT 215
DB 175 KANITVTLHDIIHTKKTLLFVFEVYLDKDLKQYDDCGNVTNMKNVTLFLFOLLRGLAYC 233
QY 216 HHQVHLHRDLKPOLNLTISHGELKLDADFGIARAKSIPSQTSSESVTLWYRPDALLGAT 275
DB 234 HHRKTLHRDLKPOLNLTINRGELKLDADFGIARAKSIPTKTYSNVVTLWYRPDVLGSGT 293
QY 276 EYSESLDIWAGCIFIEMFGQPLPFGVSNILEOLEKIWEVLGVPTEDTWPVGSKLPNYN 335
DB 294 EYSTDIMWGVCIFYEMATGRPLFPFG-STVKEELHLIFRLLGTPTEDTWPGLISNEEFPRT 352
QY 336 PEWFP--LPTPRSLHVMNRIGRVPFAEDLASOMLKGFPRDRVSQGLVHDFYSLALPQ 393
DB 353 YNPYKTRABAL-----LHAPRLBECGDADLTLYLLQFEGRNRIISADNAKHPFLSLGB 408

```

```

QY 394 LYQLPDEESLFTVSGVRLKPE 414
DB 409 VHQLDHTASIFSILKEIQLQKD 429

RESULT 6
S32384
protein kinase (EC 2.7.1.37) cdc2-related PCTAIR-2 - human
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-2004
C:Accession: S32384; S22746
R:Heyerson, M.; Ender, G.H.; Wu, C.L.; Su, L.K.; Gorke, C.; Nelson, C.; Harlow, E.; Tsai
EMBO J. 11, 2909-2917, 1992
A>Title: A family of human cdc2-related protein kinases.
A:Reference number: S23382; MUID:92347325; PMID:1639063
A:Accession: S23384
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-523 <MEY>
A:Cross-references: UNIPROT:Q00537; EMBL:X66360
R:Meyerson, M.L.
submitted to the EMBL Data Library, May 1992
A:Reference number: S22743
A:Accession: S22746
A:Molecule type: mRNA
A:Residues: 1-366, 'L', 368-432, 'E', 434-523 <ME2>
A:Cross-references: EMBL:X66360; NID:936616; PIDN:CAA47004.1; PID:936617
C:Superfamily: protein kinase homology
C:Keyword: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:190-443/Domain: protein kinase homology <KIN>
F:198-206/Region: protein kinase ATP-binding motif

Query Match 38.7%; Score 897; DB 2; Length 523;
Best Local Similarity 52.4%; Pred. No. 4.4e-34;
Matches 177; Conservative 62; Mismatches 81; Indels 18; Gaps 6;

QY 97 FGAASYLNLKLEGGSYATVYKGISRINGQVALKVISMNAEAGVPFPAIRBASLLKGL 156
DB 186 FPKMEYIKLEKLEGGSYATVYKGSKLTIDNLVALKEIRLEHEEGAPCTAIRREVSLIKDL 245
QY 157 KANITVTLHDIIHTKKTLLFVFEVYMHDTLAQVMSQHPGLHPHNVRLLPFWOLLRGLAYTH 216
DB 246 KANITVTLHDIIHTKKTLLFVFEVYLDKDLKQYDDCGNVTNMKNVTLFLYOLLRGLAYCH 305
QY 217 HOHVLHRDLKPOLNLTISHGELKLDADFGIARAKSIPSQTSSESVTLWYRPDALLGATE 276
DB 306 RKVHLRDLKPOLNLTINRGELKLDADFGIARAKSIPTKTYSNVVTLWYRPDVLGSGSE 365
QY 277 YSESLDIWAGCIFIEMFGQPLPFGVSNILEOLEKIWEVLGVPTEDTWPVGS--KLPN 333
DB 366 YSQIDIMWGVCIFYEMASGRPLFPFG-STVEDELHLIFRLLGTPTEDTWPGLISNEEFPRT 424
QY 334 YNPYKTRABAL-----LHAPRLBECGDADLTLYLLQFEGRNRIISADNAKHPFLSLGB 389
DB 425 YN--FPKTKPPL-----INHAPRLBECGIELRLKFLQYSSKRVSAEBAKHYFRS 475
QY 390 LPSQLYQLPDEESLFTVSGVRLKPEWCDLLASY-OKGH 426
DB 476 LGPRTHALPESVIFSILKEIQLQKDPGRFNSSYPETGH 513

RESULT 7
S32831
serine/threonine-specific protein kinase PCTAIR-3 (EC 2.7.1.1.-) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1993 #sequence_revision 02-Jun-1994 #text_change 16-Aug-2004
C:Accession: S32831
R:Heyerson, M.; Ender, G.H.; Wu, C.L.; Su, L.K.; Gorke, C.; Nelson, C.; Harlow, E.; Tsai
EMBO J. 11, 2909-2917, 1992
A>Title: A family of human cdc2-related protein kinases.
A:Reference number: S23382; MUID:92347325; PMID:1639063
A:Accession: S32831

```

A:Molecule type: mRNA
 A:Residues: 1380 <MEY>
 A:Cross-references: UNIPROT:Q07002; EMBL:X66362; NID:g297101; PIDN:CAA47005.1; PID:g2971
 C:Superfamily: protein kinase homology
 C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 F:48-300/Domain: protein kinase homology <KIN>
 F:56-64/Region: protein kinase ATP-binding motif

Query Match 37.8%; Score 876.5; DB 2; Length 380;
 Best Local Similarity 54.2%; Pred. No. 2.8e-33;
 Matches 175; Conservative 52; Mismatches 83; Indels 13; Gaps 5;

QY 97 FGAAASYLNLEKLGSGSYATVYKGISRINGQLVALKVISMAEEGVPTAIRASLLKGL 156
 DB 44 FGLLETVYKLDKLGEGTYATVFKRSKLTENLVALKETIRHEEGAPCTAIRVSLKGL 103
 QY 157 KHANIVLHDIHTKETLTFVFEYMTDLAQYMSQHPGGL-HPHNVRLFMFOLLRLGLAYI 215
 DB 104 KHANIVLHDLHTDRLSLTFVFEYLDSDLKQYL-DHCGNLSMHNKIFMOLLRLGLAYC 162
 QY 216 HHQVTLHRDLKPNOLLISHLGELKADFGIARAKSIPSOYSSVVTLWYRPDALLGAT 275
 DB 163 HTKXILHRDLKPNOLLINERGELKADFGIARAKSVPTKYSNEVVTLWYRPDPVLLGST 222
 QY 276 EYSESLDIMGAGCIFTEMFGQGLFPGVSNILEQLEKIMEVGLVPTEDTPGVSUKLPNN 335
 DB 223 EYSTPLAMGVGCIHYEMATGRPLPFG-STVKEELHLIFLLGTPTBETWPGVTARSEFR 281
 QY 336 PEMFPLPTPRSLHVVNNRLGRVBEA---DLASQMLKGFPRDRVSAQOALVHDYFSALP 391
 DB 282 TYSFPCYLPQPL-----INHARRLDTDGIHLLSLLVESKRMABALSHSYRSISG 335
 QY 392 SOLYQLPDESLFTVSGVRLKPE 414
 DB 336 ERVHQLDPTASIFSLKEIQLQKD 358

RESULT 8

B87722
 protein ZC123.4 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 16-Aug-2004
 C:Accession: B87722
 R:Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
 A:Reference number: A75000; NCID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_eli
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: B87722
 A:status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-378 <STO>
 A:Cross-references: UNIPROT:Q45022; GB:chr.I; PIDN:AAB97606.1; PID:g2804502; GSPDB:GN000
 A:Note: contains similarity to ser/thr protein kinases
 C:Genetics:
 A:Gene: ZC123.4
 A:Map position: 1
 C:Superfamily: discoidin I amino-terminal homology; fibronectin type III repeat homology
 inase homology; SAM homology; SH2 homology; SH3 homology

Query Match 35.3%; Score 819; DB 2; Length 378;
 Best Local Similarity 51.2%; Pred. No. 1.1e-30;
 Matches 165; Conservative 53; Mismatches 92; Indels 12; Gaps 4;

QY 103 YLNLEKLGSGSYATVYKGISRINGQLVALKVISMAEEGVPTAIRASLLKGLKHANTY 162
 DB 60 YKRIDLKGGSYATVYKCGSKLDGSLVALKEIKLQCGELPFAIRASLLKRLKRLANTY 119
 QY 163 LHHDIHTKETLTFVFEYMTDLAQYMSQHPGGLHPHNVRLFMFOLLRLGLAYIHHQVLA 222
 DB 120 SLHDIYQHNLTFVFEYMTDLAQYMSQHPGGLHPHNVRLFMFOLLRLGLAYIHHQVLA 179

QY 223 RDLKPNOLLISHLGELKADFGIARAKSIPSOYSSVVTLWYRPDALLGATEYSELD 282
 DB 180 RDLKPNOLLISHLGELKADFGIARAKSVPSRTYHEVVTLWYRPDVLMSGTIDYSTLD 239
 QY 283 IMGAGCIFTEMFGQGLFPGVSN-----LLBQLEKIMEVGLVPTEDTPGVSUKLPNNYNE 337
 DB 240 MNGVGCIFAEICTGALFPGSKDSHPGTDQDMIFSRIGTPDEKKMPREVTLROYTPE 299
 QY 338 WEPPLPTPRSLHVVNNRLGRV-DEADLASQMLKGFPRDRVSAQOALVHDYFSALP 396
 DB 300 LEPFRYELSLFIAVNNPMTKILKTGOELGMLQLRPSRVSASAMLHYPFASPREVAL 359
 QY 397 LPDESLFTVSGVRLKPEKCDL 418
 DB 360 LAPQSIF-----RLK-ELKDL 375

RESULT 9

T15445
 hypothetical protein C07G1.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2004
 C:Accession: T15445
 R:Hawkins, J.
 submitted to the EMBL Data Library, May 1996
 A:Description: The sequence of C. elegans coemid C07G1.
 A:Reference number: Z18352
 A:Accession: T15445
 A:status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-577 <HAW>
 A:Cross-references: UNIPROT:Q17794; EMBL:U58751; NID:g1326379; PID:g1326380; PIDN:AAB006;
 A:Experimental source: strain Bristol N2; clone C07G1
 C:Genetics:
 A:Gene: CESP:C07G1.3
 A:Map position: 4
 A:introns: 35/1; 101/2; 136/1; 168/2; 185/3; 267/3; 311/3; 449/2; 525/3
 C:Superfamily: protein kinase homology

Query Match 34.5%; Score 800.5; DB 2; Length 577;
 Best Local Similarity 49.3%; Pred. No. 1.1e-29;
 Matches 167; Conservative 50; Mismatches 97; Indels 25; Gaps 6;

QY 97 FGAAASYLNLEKLGSGSYATVYKGISRINGQLVALKVISMAEEGVPTAIRASLLKGL 156
 DB 232 FGLLETVYKLDKLGEGTYATVFKRSKLTENLVALKETIRHEEGAPCTAIRVSLKGL 291
 QY 157 KHANIVLHDIHTKETLTFVFEYMTDLAQYMSQHPGGLHPHNVRLFMFOLLRLGLAYI 216
 DB 292 KHANIVLHDIHTDRLSLTFVFEYLDSDLKQYMDSCNNAQMNNIRLFLYQLRLGLAYC 351
 QY 217 HHQVTLHRDLKPNOLLISHLGELKADFGIARAKSIPSOYSSVVTLWYRPDALLGATE 276
 DB 352 QRVLVHRDLKPNOLLITAKGELKADFGIARAKSVPTKYSNEVVTLWYRPDVLGSTD 411
 QY 277 YSESLDIMGAGCIFTEMFGQGLFPGVSNILEQLEKIMEVGLVPTEDTPGVSUKLPNNP 336
 DB 412 YSTHIDMNGVGCIFTEMAGRALFPG-STPTEQLILFRITGSPRPDHPPTCEKPTTYP 470
 QY 337 ---EMFPLPTPRSLHVVNNRLGRVBEA---DLASQMLKGFPRDRVSAQOALVHDYFS 388
 DB 471 YANRYNDEPLCR-----QIPRIDAHGFEMLKFLQYEGGDRVSAABAVHNPFLR 520
 QY 389 ALPSQLYLPDESLFTVSGVRLKPEKCDLASVYKGNH 427
 DB 521 TIAVKGCHLRDEQSVLEADGIHIRE---LNAS---DHN 553

RESULT 10

S40021
 protein kinase (EC 2.7.1.37) cdcd homolog - slime mold (Dictyostelium discoideum)
 C:Species: Dictyostelium discoideum
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

A:Accession: S40021; S39071
R:Michaelis, C.E.; Weeks, G.
submitted to the EMBL Data Library, August 1992
A:Description: The unicellular organism Dictyostelium discoideum possesses a highly related
A:Reference number: S40021
A:Accession: S40021
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-292 <MT>
A:CROSS-references: UNIPROT:P34117, EMBL:L06052, NID:g167695, PIDN:AAA16056.1, PTD:g167695
R:Michaelis, C., Weeks, G.
Biochim. Biophys. Acta 1179, 117-124, 1993
A>Title: The isolation from a unicellular organism, Dictyostelium discoideum, of a highly related
A:Reference number: S39071, MUID:94032415, PMID:8218353
A:Accession: S39071
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-18, 'V', 20-292 <M12>
A:CROSS-references: EMBL:L00652
C:Complex: In various organisms, cdc2 has been identified as a component of the M-phase
C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP; cell cycle control; phosphoprotein; phosphotransferase; serine/threonine
F:2-254/Domain: protein kinase homology <kin>
F:10-18/Region: protein kinase ATP-binding motif
F:33,51,126,128/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 32.1%, Score 745.5; DB 2; Length 292;
Best Local Similarity 52.2%, Pred. No. 1.9e-27;
Matches 151; Conservative 49; Mismatches 84; Indels 5; Gaps 4;

Gy 103 YLNLEKLGEGSYATVYKGISHSINQVALKYISMNAE-EGVPFAIRBASILKGLKIANT 161
Db 4 YSKIERKEGEGGYGVKNRAKNREGEIYALKRIIRDSEBEGVPCTAIREISILKELKAPNI 63
Gy 162 VLIHDIHTKETLTFVEEVWHTDLAQYMSQHPSGLHFNHNRLEFMOLLGLAYTHGHVL 221
Db 64 VRLLHDVHTERKTLTVBEYLDDQLTKTLDCEGGGISKPTIKSFMYQLKGVAFCHDRVL 123
Gy 222 HRDLKPONLLISHLGEKLADFGILARAKSIPSQTYSSSEVVLTLMWRPPDALLGATEVSSEL 281
Db 124 HRDLKPONLLINRKGEKLADFGLARAFAFGIVRTYSHEVVLTLMWRPDDLMSGRKSTPI 183
Gy 282 DIMAGCIPIEMPGOGPLFPGVSNILIQLEKIMENVLGVPEDTTPGYSKLPNTNPENFPL 341
Db 184 DIMWALCIFAEFMAAGRPDPG-SGTSDQLFRIFKILTPNEESWPSITTEPEYKTD-FPV 241
Gy 342 PTPLSLHVWNRLGRVPEAEADLASOMLKGFPRDVSQAOLNVHDFYFSL 390
Db 242 HPAHQLSIVHGDE--KGLNLISKMLQYDPNQKITAAALKKHYFFGL 288

RESULT 11
OKBY85
protein kinase PHO85 (EC 2.7.1.-) - Yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein P7102.18c-a; protein YPD03ic
C:Species: *Saccharomyces cerevisiae*
C>Date: 30-Jun-1991 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C:Accession: S62043; S05853; JCI1063; PC4036
R:Dietrich, F.S.; Allen, E.; Araujo, R.; Aparicio, A.; Carpenter, J.; Cherry, J.M.; Chun, H.; Lin, A.; Lin, D.; Marathe, R.; Mittipati, S.; Namath, A.; Oefner, P.; Petel, F.X.; submitted to the EMBL Data Library, December 1995
A:Reference number: S62043
A:Accession: S62043
A:Molecule type: DNA
A:Residues: 1-302 <DIE>
A:CROSS-references: UNIPROT:Q06888, EMBL:U44030, NID:g1171408, PID:g1171426; GSPDB:GN000
R:Ilesono, Y.; Tanaka, K.; Toh-e, A.
Nucleic Acids Res. 15, 10299-10309, 1987
A>Title: Negative regulator of the PMO system in *Saccharomyces cerevisiae*: isolation and
A:Reference number: S05853; MUID:88096566; PMID:3320965
A:Accession: S05853
A:Molecule type: DNA
A:Residues: 1-95, 'A', 97-302 <UES>

A:Cross-references: EMBL:X00867; NID:g4169; PIDN:CMAA6774.1; PID:g4170
Experimental source: strain DCS

Rizhong, H.L.; Li, B.L.; Ao, S.Z.
Acta Biochim. Biophys. Sin., 24, 523-529, 1992

A>Title: Cloning and high expression of yeast PHO85 gene in *Escherichia coli*.
A:Reference number: JCI1063
A:Accession: JCI1063

A:Molecule type: DNA
A:Residues: 1-83 'N', 86-95, 'A', 97-302 <ZHO>
A:Accession: PCJ035

A:Molecule type: protein
A:Residues: 1-15 <ZH2>

A:Genetics:
C:Gene:
A:Gene: SGD:PHO85; MIPS:YPL031c
A:Cross-references: SGD:S0005952; MIPS:YPL031c
A:Map position: 16f.

C:Superfamily: kinase-related transforming protein; protein kinase homology
C:Keywords: ATP, phosphotransferase; serine/threonine-specific protein kinase
F/2-358/Domain: protein kinase homology <KIN>
F/10-18/Region: protein kinase ATP-binding motif
F/33,50,130,132/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 32.1%; Score 745; DB 1; Length 302;
Best Local Similarity 51.7%; Pred.No.2,1e-57;
Matches 152; Conservative 44; Mismatches 88; Indels 10; Gaps 4;

Oy 103 YLNLEKLSEGSAAATYYKGISLINGQLVALKVISMAEEGVPTFAIRASLLGKHANIV 162
Db 4 FKQLEKLTNGTYATAIYYKKINKTKTTGYVVALKEVDLBEGRPTRAIRSLSMKELXENIV 63

Oy 163 LLHDIHTKETITLPVEPYWMTDLAQYSQH----PGLHPHNVRLFMFOLLRLAYITH 217
Db 64 RLVDVIHTENKTTLTFPEEFMDNDLKRYMDSRTVGNTFRGELNLVKYFOMLLOGLACHE 123

Oy 218 QHVLRDLDKPOMULLSHLGELKLADFGLARAVSPQSSTSEVTLMYPDPDALIGATEY 277
Db 124 NKILHRDLKPOMULNKRGQLKDGFGLARAFGIPIVNTSSSEVTVTLMYRPAIDLVMGRITY 183

Oy 278 SSELDINGAGCIFEMIFQQPPFPFGVSNIIEOLEKIWEVLGVTEDTWPQVSKLPNNPE 337
Db 184 STSIDWSCGCILAEMIWTGPFP--TNDEOLKLIFFDIGTGNESSLMPSVTKLPKYNP 242

Oy 338 WEPLETPPSLVHWNRRLGRVP---EAEDELASGMCLKPFPRDRVAQEALLVIDFS 368
Db 243 IQGRP-PEDLRQVLOPHYTEPLDGMLMDFLHGILLQNIPDWRLSAKOKLHHPWFA 295

RESULT 12

A45091
protein kinase (EC 2.7.1.37) cdcc-related nclk - bovine
N/A:Alternate names: cdcc-related protein kinase PSSALRB homolog; rat protein kinase II late
C/Species: Bos primigenius taurus (catle)
C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_1999
C:/Accession: A45091; S39453; S39454

J.Rlew, J.J. Winkeln, R.U.: Paudel, H.K.; Wang, J.H.
U.Biol. Chem. 267, 25922-25926, 1992

A>Title: Brain proline-directed protein kinase is a neurofilament kinase which displays p
A:Reference number: A45091; MUId:93100310; PMID:1464504

A:Accession: A45091
A:Molecule type: mRNA
A:Residues: 1-292 <LEM>
A:Cross-references: GB:L04798
A:Experimental source: brain
A>Note: Sequence extracted from NCBI backbone (NCBIN:120848, NCBIP:120849)

F/R Kobayashi, S.; Ishiguro, K.; Omori, A.; Takamatsu, M.; Arioka, M.; Imahori, K.; Uchida,
FEBS Lett. 335, 171-175, 1993

A>Title: A cdcc-related kinase PSSALRE/cdks is homologous with the 30 kDa subunit of tau
A:Reference number: S39453; MUId:94074679; PMID:8253190
A:Accession: S39453
A:Molecule type: mRNA
A:Residues: 1-292 <KO>
A:Cross-references: GB:X02440; NID:9572618; PIDN:CMAA57821.1; PID:9572619
A:Accession: S39454

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2004, 12:15:39 / Search time 355 Seconds
(without alignments)
440.045 Million cell updates/sec

Title: US-10-786-065-5

Perfect score: 2320
Sequence: 1 MGOELCAKTVOGCGSCYHCS.....COLLASYCKGHHAPGSKCW 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1595201 seqs, 359116952 residues

Total number of hits satisfying chosen parameters: 1595201

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2320	100.0	435	9	US-09-801-861-5 Sequence 5, Appl1
2	2320	100.0	435	14	US-10-224-562-5 Sequence 5, Appl1
3	2320	100.0	435	17	US-10-618-941-86 Sequence 86, Appl1
4	1784	76.9	443	9	US-09-801-861-2 Sequence 2, Appl1
5	1784	76.9	321	14	US-10-224-562-2 Sequence 2, Appl1
6	1659.5	71.5	321	15	US-10-466-759-7 Sequence 7, Appl1
7	1410	60.8	268	15	US-10-425-114-54124 Sequence 54124, A
8	1195	51.5	330	9	US-09-801-861-8 Sequence 8, Appl1
9	1195	51.5	330	14	US-10-224-562-8 Sequence 8, Appl1
10	1194.5	51.5	423	17	US-10-153-242-2 Sequence 2, Appl1
11	1194.5	51.5	423	14	US-10-723-860-407 Sequence 407, App
12	1191.5	51.4	423	14	US-10-153-242-4 Sequence 4, Appl1
13	1191	51.3	330	9	US-09-801-861-9 Sequence 9, Appl1

14	1191	51.3	330	14	US-10-224-562-9	Sequence 9, Appl1
15	1190	51.3	330	9	US-09-801-861-10	Sequence 10, Appl1
16	1190	51.3	330	14	US-10-224-562-10	Sequence 10, Appl1
17	1185.5	51.1	469	16	US-10-408-765A-1291	Sequence 1291, Ap
18	1000	43.1	245	9	US-09-801-861-7	Sequence 7, Appl1
19	1000	43.1	245	14	US-10-224-562-7	Sequence 7, Appl1
20	993	42.8	240	9	US-09-801-861-6	Sequence 6, Appl1
21	993	42.8	240	14	US-10-224-562-6	Sequence 6, Appl1
22	916	39.5	496	14	US-10-354-358-64	Sequence 64, Appl1
23	916	39.5	496	14	US-10-354-358-64	Sequence 2263, Ap
24	910	39.2	527	9	US-09-925-298-450	Sequence 490, App
25	910	39.2	527	14	US-10-102-806-490	Sequence 490, App
26	909.5	39.2	482	15	US-10-112-944-586	Sequence 286, App
27	900	38.8	460	13	US-10-087-192-1299	Sequence 1299, Ap
28	897.5	38.7	504	17	US-10-256-106-13	Sequence 13, Appl
29	897.5	38.7	504	17	US-10-618-941-85	Sequence 85, Appl
30	897	38.7	437	15	US-10-296-115-1222	Sequence 1222, Ap
31	895	38.6	481	13	US-10-087-192-1302	Sequence 1302, Ap
32	893	38.5	472	17	US-10-851-921-8	Sequence 8, Appl1
33	890	38.4	523	14	US-10-354-358-86	Sequence 86, Appl1
34	876.5	37.8	374	14	US-10-325-430-6	Sequence 6, Appl1
35	876.5	37.8	380	17	US-10-757-262-50	Sequence 50, Appl
36	861	37.1	519	13	US-10-087-192-1506	Sequence 1506, Ap
37	819	35.3	378	14	US-10-369-493-4977	Sequence 4977, Ap
38	800.5	34.5	577	14	US-10-369-493-6067	Sequence 6067, Ap
39	785.5	33.9	326	15	US-10-389-566-2455	Sequence 2455, Ap
40	763	32.9	334	14	US-10-389-493-3434	Sequence 3434, Ap
41	759	32.7	366	15	US-10-389-566-2114	Sequence 2114, Ap
42	752	32.4	320	15	US-10-389-566-2115	Sequence 2115, Ap
43	751	32.4	305	9	US-09-801-368-262	Sequence 262, App
44	751	32.4	305	15	US-10-389-566-1838	Sequence 1838, Ap
45	751	32.4	305	15	US-10-389-566-2066	Sequence 2066, Ap

ALIGNMENTS

RESULT 1
US-09-801-861-5
; Sequence 5, Application US/09801861
; Patent No. US20020119544A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Human
US-09-801-861-5

Query Match 100.0%, Score 2320, DB 9, Length 435,
Best Local Similarity 100.0%, Pred. No. 6,2e-183,
Matches 435, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

QY	1	MGOELCAKTVOGCGSCYHCS	EGEHAHSCRRSPETTEAFAFKLTDLKEASCSMTSPHRL	60
DB	1	MGOELCAKTVOGCGSCYHCS	EGEHAHSCRRSPETTEAFAFKLTDLKEASCSMTSPHRL	60
QY	61	QAAAROKFKSKPRNSDPCFOEEDRROGFOMKSLPFGAASSYLNLKLGEGSYATVYKG	120	
DB	61	QAAAROKFKSKPRNSDPCFOEEDRROGFOMKSLPFGAASSYLNLKLGEGSYATVYKG	120	
QY	121	ISIRINGOLVALKVISMNAEGVPTFAIRASLILKGLKIANIVLHDIIHTKETTFLVFEX	180	
DB	121	ISIRINGOLVALKVISMNAEGVPTFAIRASLILKGLKIANIVLHDIIHTKETTFLVFEX	180	

QY 181 MHTDLAQYMSQHFGGLHPHNVRLFMFQLRLGLAYIHHQVHLRDLKPNQMLLSHLGELKL 240
DB 181 MHTDLAQYMSQHFGGLHPHNVRLFMFQLRLGLAYIHHQVHLRDLKPNQMLLSHLGELKL 240
QY 241 ADFGLARAKSIPQYTSSEVVTLMYRPPDALLGATEYSELDIWGAGCIFIEMFOGQPLF 300
DB 241 ADFGLARAKSIPQYTSSEVVTLMYRPPDALLGATEYSELDIWGAGCIFIEMFOGQPLF 300
QY 301 PGVSNILBQLEKIMEVLGVPTEDTWPQVSKLPYNNPEWFPPLPTPRSLHYVWNLGRVPEA 360
DB 301 PGVSNILBQLEKIMEVLGVPTEDTWPQVSKLPYNNPEWFPPLPTPRSLHYVWNLGRVPEA 360
QY 361 EDLASQMLKGFPRDRVSAQDALVHDYFSAALPSQLYQLPDEBSLFTVSGVRLKPEMCDLLA 420
DB 361 EDLASQMLKGFPRDRVSAQDALVHDYFSAALPSQLYQLPDEBSLFTVSGVRLKPEMCDLLA 420
QY 421 SYOKGHHPAQFSKCM 435
DB 421 SYOKGHHPAQFSKCM 435

RESULT 2

US-10-224-562-5
; Sequence 5, Application US/10224562
; Publication No. US2003002229A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001098DIV
; CURRENT APPLICATION NUMBER: US/10/224,562
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-224-562-5

Query Match 100.0%; Score 2320; DB 14; Length 435;
Best Local Similarity 100.0%; Pred. No. 6,2e-183;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGEELCAKTVQPGCSCHYCSGEGEHAHSCRRSQPETTEAFAFKLTDLKEASCMTSFPHPRL 60
DB 1 MGEELCAKTVQPGCSCHYCSGEGEHAHSCRRSQPETTEAFAFKLTDLKEASCMTSFPHPRL 60
QY 61 QARARQKFKSKRPRNSDCFOEEDLRQGFQWRKSLPFGAASLYLNLEKLGEGSYATVYKG 120
DB 61 QARARQKFKSKRPRNSDCFOEEDLRQGFQWRKSLPFGAASLYLNLEKLGEGSYATVYKG 120
QY 121 ISRINGQVLAALKVISMNAEBGVFTAIRBASLLKGLKHANIVLHDIHTKETLTFVFEY 180
DB 121 ISRINGQVLAALKVISMNAEBGVFTAIRBASLLKGLKHANIVLHDIHTKETLTFVFEY 180
QY 181 MHTDLAQYMSQHFGGLHPHNVRLFMFQLRLGLAYIHHQVHLRDLKPNQMLLSHLGELKL 240
DB 181 MHTDLAQYMSQHFGGLHPHNVRLFMFQLRLGLAYIHHQVHLRDLKPNQMLLSHLGELKL 240
QY 241 ADFGLARAKSIPQYTSSEVVTLMYRPPDALLGATEYSELDIWGAGCIFIEMFOGQPLF 300
DB 241 ADFGLARAKSIPQYTSSEVVTLMYRPPDALLGATEYSELDIWGAGCIFIEMFOGQPLF 300
QY 301 PGVSNILBQLEKIMEVLGVPTEDTWPQVSKLPYNNPEWFPPLPTPRSLHYVWNLGRVPEA 360
DB 301 PGVSNILBQLEKIMEVLGVPTEDTWPQVSKLPYNNPEWFPPLPTPRSLHYVWNLGRVPEA 360
QY 361 EDLASQMLKGFPRDRVSAQDALVHDYFSAALPSQLYQLPDEBSLFTVSGVRLKPEMCDLLA 420
DB 361 EDLASQMLKGFPRDRVSAQDALVHDYFSAALPSQLYQLPDEBSLFTVSGVRLKPEMCDLLA 420

QY 421 SYOKGHHPAQFSKCM 435
DB 421 SYOKGHHPAQFSKCM 435

RESULT 3

US-10-618-941-86
; Sequence 86, Application US/10618941
; Publication No. US2004019792A1
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 86
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-941-86

Query Match 100.0%; Score 2320; DB 17; Length 435;
Best Local Similarity 100.0%; Pred. No. 6,2e-183;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGEELCAKTVQPGCSCHYCSGEGEHAHSCRRSQPETTEAFAFKLTDLKEASCMTSFPHPRL 60
DB 1 MGEELCAKTVQPGCSCHYCSGEGEHAHSCRRSQPETTEAFAFKLTDLKEASCMTSFPHPRL 60
QY 61 QARARQKFKSKRPRNSDCFOEEDLRQGFQWRKSLPFGAASLYLNLEKLGEGSYATVYKG 120
DB 61 QARARQKFKSKRPRNSDCFOEEDLRQGFQWRKSLPFGAASLYLNLEKLGEGSYATVYKG 120
QY 121 ISRINGQVLAALKVISMNAEBGVFTAIRBASLLKGLKHANIVLHDIHTKETLTFVFEY 180
DB 121 ISRINGQVLAALKVISMNAEBGVFTAIRBASLLKGLKHANIVLHDIHTKETLTFVFEY 180
QY 181 MHTDLAQYMSQHFGGLHPHNVRLFMFQLRLGLAYIHHQVHLRDLKPNQMLLSHLGELKL 240
DB 181 MHTDLAQYMSQHFGGLHPHNVRLFMFQLRLGLAYIHHQVHLRDLKPNQMLLSHLGELKL 240
QY 241 ADFGLARAKSIPQYTSSEVVTLMYRPPDALLGATEYSELDIWGAGCIFIEMFOGQPLF 300
DB 241 ADFGLARAKSIPQYTSSEVVTLMYRPPDALLGATEYSELDIWGAGCIFIEMFOGQPLF 300
QY 301 PGVSNILBQLEKIMEVLGVPTEDTWPQVSKLPYNNPEWFPPLPTPRSLHYVWNLGRVPEA 360
DB 301 PGVSNILBQLEKIMEVLGVPTEDTWPQVSKLPYNNPEWFPPLPTPRSLHYVWNLGRVPEA 360
QY 361 EDLASQMLKGFPRDRVSAQDALVHDYFSAALPSQLYQLPDEBSLFTVSGVRLKPEMCDLLA 420
DB 361 EDLASQMLKGFPRDRVSAQDALVHDYFSAALPSQLYQLPDEBSLFTVSGVRLKPEMCDLLA 420
QY 421 SYOKGHHPAQFSKCM 435
DB 421 SYOKGHHPAQFSKCM 435

RESULT 4

US-09-801-861-2
; Sequence 2, Application US/09801861
; Patent No. US20020119544A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: THEROP

FILE REFERENCE: CL001098
CURRENT APPLICATION NUMBER: US/09/801,861
CURRENT FILING DATE: 2001-03-09
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 343
TYPE: PRT
ORGANISM: Human
US-09-801-861-2

Query Match 76.9%; Score 1784; DB 9; Length 343;
Best Local Similarity 100.0%; Pred. No. 8,4e-139; Indels 0; Gaps 0;
Matches 336; Conservative 0; Mismatches 0;

QY 1 MGQELCAKTVQPCSCYHSGEGEAHSCRROPEETEAFLKLDLEASCSMTSPHPRGL 60
DB 1 MGQELCAKTVQPCSCYHSGEGEAHSCRROPEETEAFLKLDLEASCSMTSPHPRGL 60
QY 61 QARARQKFKSKPRNSDCFOEDLROGFQWRKSLPFGAASLYLNLEKLGEGSYATVYKG 120
DB 61 QARARQKFKSKPRNSDCFOEDLROGFQWRKSLPFGAASLYLNLEKLGEGSYATVYKG 120
QY 121 ISRINGQVALKVIISMAEEGVFTAIRASLLKGLKHANIVLHDIHTKETLTFFVEX 180
DB 121 ISRINGQVALKVIISMAEEGVFTAIRASLLKGLKHANIVLHDIHTKETLTFFVEX 180
QY 181 MHTDLAQVMSQHGGHHPNVRLFMFQDLRGLAYIHGHVLRDLKPOLLLSHLGELKL 240
DB 181 MHTDLAQVMSQHGGHHPNVRLFMFQDLRGLAYIHGHVLRDLKPOLLLSHLGELKL 240
QY 241 ADFGLARAKSIPQYSSSEVVTLMWRPPDALLGATEYSSELDIWAGCIFIEMFOGQPLF 300
DB 241 ADFGLARAKSIPQYSSSEVVTLMWRPPDALLGATEYSSELDIWAGCIFIEMFOGQPLF 300
QY 301 PGVSNLEOLEKIMEVLGVPTEDTPGVSCLKPNYP 336
DB 301 PGVSNLEOLEKIMEVLGVPTEDTPGVSCLKPNYP 336

RESULT 5
US-10-224-562-2
Sequence 2, Application US/10224562
Publication No. US20030022229A1
GENERAL INFORMATION:
APPLICANT: YAN, Chunhua et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: CL001098DI
CURRENT APPLICATION NUMBER: US/10/224,562
CURRENT FILING DATE: 2002-08-21
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 343
TYPE: PRT
ORGANISM: Homo sapiens
US-10-224-562-2

Query Match 76.9%; Score 1784; DB 14; Length 343;
Best Local Similarity 100.0%; Pred. No. 8,4e-139; Indels 0; Gaps 0;
Matches 336; Conservative 0; Mismatches 0;

QY 1 MGQELCAKTVQPCSCYHSGEGEAHSCRROPEETEAFLKLDLEASCSMTSPHPRGL 60
DB 1 MGQELCAKTVQPCSCYHSGEGEAHSCRROPEETEAFLKLDLEASCSMTSPHPRGL 60
QY 61 QARARQKFKSKPRNSDCFOEDLROGFQWRKSLPFGAASLYLNLEKLGEGSYATVYKG 120
DB 61 QARARQKFKSKPRNSDCFOEDLROGFQWRKSLPFGAASLYLNLEKLGEGSYATVYKG 120
QY 121 ISRINGQVALKVIISMAEEGVFTAIRASLLKGLKHANIVLHDIHTKETLTFFVEX 180

DB 121 ISRINGQVALKVIISMAEEGVFTAIRASLLKGLKHANIVLHDIHTKETLTFFVEX 180
QY 181 MHTDLAQVMSQHGGHHPNVRLFMFQDLRGLAYIHGHVLRDLKPOLLLSHLGELKL 240
DB 181 MHTDLAQVMSQHGGHHPNVRLFMFQDLRGLAYIHGHVLRDLKPOLLLSHLGELKL 240
QY 241 ADFGLARAKSIPQYSSSEVVTLMWRPPDALLGATEYSSELDIWAGCIFIEMFOGQPLF 300
DB 241 ADFGLARAKSIPQYSSSEVVTLMWRPPDALLGATEYSSELDIWAGCIFIEMFOGQPLF 300
QY 301 PGVSNLEOLEKIMEVLGVPTEDTPGVSCLKPNYP 336
DB 301 PGVSNLEOLEKIMEVLGVPTEDTPGVSCLKPNYP 336

RESULT 6
US-10-466-759-7
Sequence 7, Application US/10466759
Publication No. US20040081983A1
GENERAL INFORMATION:
APPLICANT: INCYTE CORPORATION
APPLICANT: LEE, Ernestine A.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: ISON, Craig H.
APPLICANT: GURURAJAN, Rajagopal
APPLICANT: ARVIZU, Chandra S.
APPLICANT: YAO, Monique G.
APPLICANT: JACKSON, Jennifer L.
APPLICANT: TANG, Y. Tom
APPLICANT: YUE, Henry
APPLICANT: TRAN, Bao
APPLICANT: DING, Li
APPLICANT: DU, Dzung Aina M.
APPLICANT: LAL, Preeti G.
APPLICANT: WARREN, Bridget A.
TITLE OF INVENTION: KINASES AND PHOSPHATASES
FILE REFERENCE: PI-0344 USN
CURRENT APPLICATION NUMBER: US/10/466,759
CURRENT FILING DATE: 2003-07-17
PRIORITY APPLICATION NUMBER: PCT/US02/01369
PRIORITY FILING DATE: 2002-01-16
PRIORITY APPLICATION NUMBER: US 60/263,083
PRIORITY FILING DATE: 2001-01-18
PRIORITY APPLICATION NUMBER: US 60/271,205
PRIORITY FILING DATE: 2001-02-23
PRIORITY APPLICATION NUMBER: US 60/271,117
PRIORITY FILING DATE: 2001-02-23
PRIORITY APPLICATION NUMBER: US 60/276,859
PRIORITY FILING DATE: 2001-03-16
PRIORITY APPLICATION NUMBER: US 60/278,504
PRIORITY FILING DATE: 2001-03-23
PRIORITY APPLICATION NUMBER: US 60/278,522
PRIORITY FILING DATE: 2001-03-23
PRIORITY APPLICATION NUMBER: US 60/280,510
PRIORITY FILING DATE: 2001-03-29
PRIORITY APPLICATION NUMBER: US 60/280,266
PRIORITY FILING DATE: 2001-03-29
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PERL Program
SEQ ID NO 7
LENGTH: 321
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE: NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: 7494145CD1
US-10-466-759-7

Query Match 71.5%; Score 1659.5; DB 15; Length 321;
Best Local Similarity 83.6%; Pred. No. 1.5e-126; Indels 63; Gaps 1;
Matches 321; Conservative 0; Mismatches 0;

```

QY 52 MTSFHRGLOAARAOXFKSKRRPSNDCFOEBDLRQGFQWRKSLPFGAASSTYINLEKGE 111
DB 1 MTSFHRGLOAARAOXFKSKRRPSNDCFOEBDLRQGFQWRKSLPFGAASSTYINLEKGE 60
QY 112 GSATATYKGSIRINGQVLVAKVISMAABEGVPTALIREASILKGLKHANIVLHDIHTK 171
DB 61 GSATATYKGSIRINGQVLVAKVISMAABEGVPTALIREASILKGLKHANIVLHDIHTK 120
QY 172 ETLTFVEVYHTDLAQYMSQHPGLHPHNVRLEFMQGLRAYIHQVHLHRDLKPQNL 231
DB 121 ETLTFVEVYHTDLAQYMSQHPGLHPHNVRLEFMQGLRAYIHQVHLHRDLKPQNL 180
QY 232 ISHIGELKADFGIARAKSIPSQTSSEVVTLMYRPDALLGATYESSLEIDTAGACITF 291
DB 181 ISHIGELKADFGIARAKSIPSQTSSEVVTLMYRPDALLGATYESSLEIDTAGACITF 240
QY 292 EMFOGQPLPFGVSNIIIEOLEKIMEVLGVPTEDTPGVSKLPYNNPWFPLPTPRSLHYV 351
DB 241 EMFOGQPLPFGVSNIIIEOLEKIMEVLGVPTEDTPGVSKLPYNNPWFPLPTPRSLHYV 285
QY 352 NLRGRVPEADLASQMLKGFPRDRVSAQBALVHDYFSAIPSOYLQDPDESLFTVSGVRL 411
DB 286 -----ESLFTVSGVRL 297
QY 412 KPDMCDLLASYOKGHHPAQFSKCM 435
DB 298 KPDMCDLLASYOKGHHPAQFSKCM 321

```

RESULT 7

```

US-10-425-114-54124
; Sequence 34124, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven B.
; APPLICANT: Tabaska, Jack B.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54124
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4119-079-Cl_Flt_Pep
US-10-425-114-54124

```

Query Match 60.8%; Score 1410; DB 15; Length 268;

Best Local Similarity 100.0%; Pred. No. 4.8e-108; Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 150 ASLLKGLKHANIVLHDIHTKETTTFVFEYMATDLAQYMSQHPGLHPHNVRLEFMQGL 209
DB 2 ASLLKGLKHANIVLHDIHTKETTTFVFEYMATDLAQYMSQHPGLHPHNVRLEFMQGL 61
QY 210 RGLAYIHQVHLHRDLKPQNLISHIGELKADFGIARAKSIPSQTSSEVVTLMYRPD 269
DB 62 RGLAYIHQVHLHRDLKPQNLISHIGELKADFGIARAKSIPSQTSSEVVTLMYRPD 121
QY 270 ALGATYESSLEIDTAGACITFEMFOGQPLPFGVSNIIIEOLEKIMEVLGVPTEDTPGV 329
DB 122 ALGATYESSLEIDTAGACITFEMFOGQPLPFGVSNIIIEOLEKIMEVLGVPTEDTPGV 181
QY 330 KLPYNNPWFPLPTPRSLHYVNRNLGRVPEADLASQMLKGFPRDRVSAQBALVHDYFSA 389

```

```

DB 182 KLPYNNPWFPLPTPRSLHYVNRNLGRVPEADLASQMLKGFPRDRVSAQBALVHDYFSA 241
QY 390 LPSQLYQLPDEESLFTVSGVRLKPE 414
DB 242 LPSQLYQLPDEESLFTVSGVRLKPE 266

```

RESULT 8

```

US-09-801-861-8
; Sequence 8, Application US/09801861
; Patent No. US20020119544A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-801-861-8

```

Query Match 51.5%; Score 1195; DB 9; Length 330;

Best Local Similarity 68.2%; Pred. No. 3.6e-90; Matches 225; Conservative 45; Mismatches 60; Indels 0; Gaps 0;

```

QY 97 FGAASSYINLEKLGSGYATVYKGSIRINGQVLVAKVISMAABEGVPTALIREASILKGL 156
DB 1 FGAASSYINLEKLGSGYATVYKGSIRINGQVLVAKVISMAABEGVPTALIREASILKGL 60
QY 157 KHANIVLHDIHTKETTTFVFEYMATDLAQYMSQHPGLHPHNVRLEFMQGLRAYIH 216
DB 61 KHANIVLHDIHTKETTTFVFEYMATDLAQYMSQHPGLHPHNVRLEFMQGLRAYIH 120
QY 217 HOHVLHRDLKPQNLISHIGELKADFGIARAKSIPSQTSSEVVTLMYRPDALLGATE 276
DB 121 HOHVLHRDLKPQNLISHIGELKADFGIARAKSIPSQTSSEVVTLMYRPDALLGATE 180
QY 277 YSELDIWDAGCITFEMFOGQPLPFGVSNIIIEOLEKIMEVLGVPTEDTPGVSKLPYNNP 336
DB 181 YSELDIWDAGCITFEMFOGQPLPFGVSNIIIEOLEKIMEVLGVPTEDTPGVSKLPYNNP 240
QY 337 EMFPLPTPRSLHYVNRNLGRVPEADLASQMLKGFPRDRVSAQBALVHDYFSAIPSOYLQ 396
DB 241 EMFPLPTPRSLHYVNRNLGRVPEADLASQMLKGFPRDRVSAQBALVHDYFSAIPSOYLQ 300
QY 397 LPDEESLFTVSGVRLKPEKCDLLASYOKGH 426
DB 301 LPDEESLFTVSGVRLKPEKCDLLASYOKGH 330

```

RESULT 9

```

US-10-224-562-8
; Sequence 8, Application US/10224562
; Publication No. US20030022229A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001098DIV
; CURRENT APPLICATION NUMBER: US/10/224,562
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 330
; TYPE: PRT

```

ORGANISM: Mus musculus
US-10-224-562-8

Query Match 51.5%; Score 1195; DB 14; Length 330;
Best Local Similarity 68.2%; Pred. No. 3.6e-90;
Matches 225; Conservative 45; Mismatches 60; Indels 0; Gaps 0;

QY 97 PBAASSYLNLEKLGEGSYATVTKGSRINGQVYALKVISMABEGVPFPAIRBASLLKGL 156
DB 1 FPKADSYELKLGEGSYATVTKGSKVNGKLVALVIRLOEBEGPFPFAIRBASLLKGL 60
QY 157 KHANIVLHDIHTKKTLLFVEFYVMTDLAQMYSQHPGGLHPNVRFLPMOLLRLGLAYIH 216
DB 61 KHANIVLHDIHTKKTLLFVEFYVMTDLAQMYSQHPGGLHPNVRFLPMOLLRLGLAYIH 120
QY 217 HOHVLHDLKPNLISHGELKADFGGLARAKSIPSTYSSSEVTLWYRPPDALLGATE 276
DB 121 QRYIHLHDLKPNLISHGELKADFGGLARAKSVSHYTSNEVTLWYRPPDALLGATE 180
QY 277 YSELDIAGAGCIFIEMFGQPLPGVSNILEQLEKIMVYLGVPEDTWPVGSKLPNVNP 336
DB 181 YSTCLDMVGICIFVEMIGVNAFPGMKIDIODLERIFVLGTPNEDTWPVGHSLPHFRP 240
QY 337 EMFPLTPRSLHVMNRIGRVEAEADLASOMLKGPRDRVSAQEAIVHDYFSALPSQLXQ 396
DB 241 ERFVYSSKSLROAMKLSYVNHAEADLASKLQCSFKNRLSAQALSHYFSDLPRLWE 300
QY 397 LPDESLFTVSGVRLKPEMCDLLASYOKGH 426
DB 301 LTMGSIFTPVNRVLRQPEAGESMRARFGKN 330

RESULT 10

US-10-153-242-2
Sequence 2, Application US/10153242
Publication No. US20030166217A1
GENERAL INFORMATION:
APPLICANT: Reinhard, Christoph
APPLICANT: Pot, David
APPLICANT: Kassam, Altaf
APPLICANT: Marenbach, Taasha
APPLICANT: Williams, Lewis T.
TITLE OF INVENTION: HUMAN CYCLIN-DEPENDENT KINASE (hPPTAIRE)
FILE REFERENCE: PP-01429.103/200130.44501
CURRENT FILING DATE: 2002-05-22
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 423
TYPE: PRT
ORGANISM: Homo sapiens
US-10-153-242-2

Query Match 51.5%; Score 1194.5; DB 14; Length 423;
Best Local Similarity 60.5%; Pred. No. 5.5e-90;
Matches 234; Conservative 55; Mismatches 83; Indels 15; Gaps 2;

QY 51 SMTSPHPRGLQARAKQFKSKPRNSDCFOEEDLRQGFQWRK-----SLPFGA 99
DB 30 TOSTIDPFEKPRANQVRVHSE---NNACINFTSTGKSPKVRHSSPSSPTSPKFK 85
QY 100 ASSYLNLEKLGEGSYATVTKGSRINGQVYALKVISMABEGVPFPAIRBASLLKGLKIA 159
DB 86 ADSYELKLGEGSYATVTKGSRVNGKLVALKVIRLOEBEGPFPFAIRBASLLKGLKIA 145
QY 160 NVLHDIHTKKTLLFVEFYVMTDLAQMYSQHPGGLHPNVRFLPMOLLRLGLAYIHQY 219
DB 146 NVLHDIHTKKTLLFVEFYVMTDLAQMYSQHPGGLHPNVRFLPMOLLRLGLAYIHQY 205
QY 220 VHLHDLKPNLISHGELKADFGGLARAKSIPSTYSSSEVTLWYRPPDALLGATEYS 279
DB 206 IHLHDLKPNLISHGELKADFGGLARAKSVSHYTSNEVTLWYRPPDALLGATEYST 265

QY 280 ELDIAGAGCIFIEMFGQPLPGVSNILEQLEKIMVYLGVPEDTWPVGSKLPNVNPWF 339
DB 266 CLDMVGICIFVEMIGVNAFPGMKIDIODLERIFVLGTPNEDTWPVGHSLPHFRP 325
QY 340 PLPTPRSLHVMNRIGRVEAEADLASOMLKGPRDRVSAQEAIVHDYFSALPSQLYQLPD 399
DB 326 TLYSSKNLRQAMKLSYVNHAEADLASKLQCSFKNRLSAQALSHYFSDLPRLMELTD 385
QY 400 EESLFTVSGVRLKPEMCDLLASYOKGH 426
DB 386 MSIFTPVNRVLRQPEAGESMRARFGKN 412

RESULT 11

US-10-723-860-407
Sequence 407, Application US/10723860
Publication No. US20040253606A1
GENERAL INFORMATION:
APPLICANT: Aziz, Nataasha
APPLICANT: Ginsburg, Wendy M.
APPLICANT: Zlocnik, Albert
TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
FILE REFERENCE: 05482.0193.NPUS01
CURRENT APPLICATION NUMBER: US/10/723, 860
CURRENT FILING DATE: 2003-11-26
PRIOR APPLICATION NUMBER: 60/429, 739
PRIOR FILING DATE: 2002-11-26
NUMBER OF SEQ ID NOS: 8393
SOFTWARE: PatentIn version 3.2
SEQ ID NO 407
LENGTH: 451
TYPE: PRT
ORGANISM: Homo sapiens
US-10-723-860-407

Query Match 51.5%; Score 1194.5; DB 17; Length 451;
Best Local Similarity 54.3%; Pred. No. 5.9e-90;
Matches 242; Conservative 61; Mismatches 106; Indels 37; Gaps 5;

QY 6 CAKTQPGGSCY-----HCEGEGBAHSCRRSQ-----PETTEAPKLTDLKEASCS 51
DB 7 CNAABPGYSAFVGTPQICVTYKSTRNCGMDSVYKPLDTTEDEKVRVQRT----- 58
QY 52 MTSFPHPRGLQARAKQFKSKPRNSDCFOEEDLRQGFQWRK-----SLPFGA 100
DB 59 OSTIDPFEKPRANQVRVHSE---NNACINFTSTGKSPKVRHSSPSSPTSPKFK 114
QY 101 SSYLNLEKLGEGSYATVTKGSRINGQVYALKVISMABEGVPFPAIRBASLLKGLKIAN 160
DB 115 DSYELKLGEGSYATVTKGSKVNGKLVALKVIRLOEBEGPFPFAIRBASLLKGLKIAN 174
QY 161 IYVLDIHTKKTLLFVEFYVMTDLAQMYSQHPGGLHPNVRFLPMOLLRLGLAYIHQY 220
DB 175 IYVLDIHTKKTLLFVEFYVMTDLAQMYSQHPGGLHPNVRFLPMOLLRLGLAYIHQY 234
QY 221 IHLHDLKPNLISHGELKADFGGLARAKSIPSTYSSSEVTLWYRPPDALLGATEYS 280
DB 235 IHLHDLKPNLISHGELKADFGGLARAKSVSHYTSNEVTLWYRPPDALLGATEYST 294
QY 281 IDIAGAGCIFIEMFGQPLPGVSNILEQLEKIMVYLGVPEDTWPVGSKLPNVNPWF 340
DB 295 IDIAGAGCIFIEMIGVNAFPGMKIDIODLERIFVLGTPNEDTWPVGHSLPHFRPRT 354
QY 341 LPTPRSLHVMNRIGRVEAEADLASOMLKGPRDRVSAQEAIVHDYFSALPSQLYQLPDE 400
DB 355 LYSKSLROAMKLSYVNHAEADLASKLQCSFKNRLSAQALSHYFSDLPRLMELTD 414
QY 401 ESLFTVSGVRLKPEMCDLLASYOKGH 426
DB 415 SSIFTPVNRVLRQPEAGESMRARFGKN 440

```

RESULT 12
US-10-153-242-4
; Sequence 4, Application US/10153242
; Publication No. US20030166217A1
; GENERAL INFORMATION:
; APPLICANT: Reinhard, Christoph
; APPLICANT: Pot, David
; APPLICANT: Kassam, Altaf
; APPLICANT: Marenbach, Tasha
; APPLICANT: Williams, Lewis T.
; TITLE OF INVENTION: HUMAN CYCLIN-DEPENDENT KINASE (hPPTAIRE)
; FILE REFERENCE: PP-01429.103/200130.445D1
; CURRENT APPLICATION NUMBER: US/10/153,242
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 423
; TYPE: PRT
; ORGANISM: mouse
US-10-153-242-4

Query Match
Best Local Similarity 51.4%; Score 1191.5; DB 14; Length 423;
Matches 233; Conservative 55; Mismatches 84; Indels 15; Gaps 2;

QY 51 SMTSPHPRGLQAARAKFKSKRPNSDCEQEDRLRQGFQWRK-----SLPFGA 99
DB 30 TSTFPPEKPAQVQRVHSE---NNACINFKSSAGKESPKVRRHSSPSPTSPKFGK 85
QY 100 ASSYLANLEKLGSGSYATVYKGISRINGQIVALKVISMNAEEGVPFAIRBASILKGLKA 159
DB 86 ASYERLEKLGSGSYATVYKGISKNGKVALKVRLOEEEGPFAIRBASILKGLKA 145
QY 160 NIVLHDIHTKETTLLFVEEYMTDLAQYWSQHPGGLHPHNVLFWFOLLRGLAYITHOH 219
DB 146 NIVLHDIHTKETTLLFVEEYVHTDLCQYMDKHPGGLHPDNVTLFLFOLLRGLSYTHQY 205
QY 220 VLRDLKPNQNLISHGLKDLADFGIARAKSIPQSYSSSEVVTLMYRPDPALLGATEYSS 279
DB 206 IHRDLKPNQNLISDGLKDLADFGIARAKSVSHSYSEVVTLMYRPDPVLLGSTEYST 265
QY 280 ELDIWAGGCIETEMFGQPLFPGVSNILBQLEKIMEVLGVPTEDTPGVSXKLPNNPEMP 339
DB 266 CLDMGVGCIETEMIGVAAFPQMKDIOQLERIFLVLTGPNEDTWPGVHSLPHFKRP 325
QY 340 PLPTPRSLHVNNRLGRVPEADLASQMLKGPFRDRVSAQELVHDYFSLPSQLYQLPD 399
DB 326 TVYNSKSLQANWKLSYVNHADLASKTLQCSFKRRLSAQALSHYFSDLPRLMLWELTD 385
QY 400 EBSLFTVSGVRLKPEKCDLLASQYQKH 426
DB 386 MSISFTVPMVRLOPEAGESMRAFGKNN 412

RESULT 13
US-09-801-861-9
; Sequence 9, Application US/09801861
; Patent No. US20020119544A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEROBOP
; FILE REFERENCE: CL001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 330
; TYPE: PRT

```

```

; ORGANISM: Human
US-09-801-861-9

Query Match
Best Local Similarity 51.3%; Score 1191; DB 9; Length 330;
Matches 224; Conservative 46; Mismatches 60; Indels 0; Gaps 0;

QY 97 FGAASSYLANLEKLGSGSYATVYKGISRINGQIVALKVISMNAEEGVPFAIRBASILKGL 156
DB 1 FGAADSYERLEKLGSGSYATVYKGISKNGKVALKVRLOEEEGPFAIRBASILKGL 60
QY 157 KANIVLHDIHTKETTLLFVEEYMTDLAQYWSQHPGGLHPHNVLFWFOLLRGLAYITH 216
DB 61 KANIVLHDIHTKETTLLFVEEYVHTDLCQYMDKHPGGLHPDNVTLFLFOLLRGLSYTH 120
QY 217 HQVLRDLKPNQNLISHGLKDLADFGIARAKSIPQSYSSSEVVTLMYRPDPALLGATE 276
DB 121 QRYIHRDLKPNQNLISDGLKDLADFGIARAKSVSHSYSEVVTLMYRPDPVLLGSTE 180
QY 277 YSEEDINGAGCIEIEMFGQPLFPGVSNILBQLEKIMEVLGVPTEDTPGVSXKLPNNP 336
DB 181 YSTCLDMGVGCIETEMIGVAAFPQMKDIOQLERIFLVLTGPNEDTWPGVHSLPHFKRP 240
QY 337 EWFPLPTPRSLHVNNRLGRVPEADLASQMLKGPFRDRVSAQELVHDYFSLPSQLYQ 396
DB 241 ERFITYSSKSLQANWKLSYVNHADLASKTLQCSFKRRLSAQALSHYFSDLPRLMLW 300
QY 397 LPDEBSLFTVSGVRLKPEKCDLLASQYQKH 426
DB 301 LTMDSISFTVPMVRLOPEAGESMRAFGKNN 330

```

```

RESULT 14
US-10-224-562-9
; Sequence 9, Application US/10224562
; Publication No. US2003002229A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001098DIV
; CURRENT APPLICATION NUMBER: US/10/224,562
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-224-562-9

Query Match
Best Local Similarity 51.3%; Score 1191; DB 14; Length 330;
Matches 224; Conservative 46; Mismatches 60; Indels 0; Gaps 0;

QY 97 FGAASSYLANLEKLGSGSYATVYKGISRINGQIVALKVISMNAEEGVPFAIRBASILKGL 156
DB 1 FGAADSYERLEKLGSGSYATVYKGISKNGKVALKVRLOEEEGPFAIRBASILKGL 60
QY 157 KANIVLHDIHTKETTLLFVEEYMTDLAQYWSQHPGGLHPHNVLFWFOLLRGLAYITH 216
DB 61 KANIVLHDIHTKETTLLFVEEYVHTDLCQYMDKHPGGLHPDNVTLFLFOLLRGLSYTH 120
QY 217 HQVLRDLKPNQNLISHGLKDLADFGIARAKSIPQSYSSSEVVTLMYRPDPALLGATE 276
DB 121 QRYIHRDLKPNQNLISDGLKDLADFGIARAKSVSHSYSEVVTLMYRPDPVLLGSTE 180
QY 277 YSEEDINGAGCIEIEMFGQPLFPGVSNILBQLEKIMEVLGVPTEDTPGVSXKLPNNP 336
DB 181 YSTCLDMGVGCIETEMIGVAAFPQMKDIOQLERIFLVLTGPNEDTWPGVHSLPHFKRP 240
QY 337 EWFPLPTPRSLHVNNRLGRVPEADLASQMLKGPFRDRVSAQELVHDYFSLPSQLYQ 396

```

Db 241 ERFLLSSSKNLQANWKLSYVNAEDLAKLQCSFKNRLSAQALSHYFSDLPRLWE 300
QY 397 LPDESLFTVSGVRLKPEKCDLLASYQKH 426
Db 301 LTDMSISFTVPVNRLOPEAGESMRAFGKN 330

RESULT 15

US-09-801-861-10
; Sequence 10, Application US/09801861
; Patent No. US20020119544A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01098
; CURRENT APPLICATION NUMBER: US/09/801,861
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-801-861-10

Query Match 51.3%; Score 1190; DB 9; Length 330;
Best Local Similarity 67.9%; Pred. No. 9, 3e-90;

Matches 224; Conservative 45; Mismatches 61; Indels 0; Gaps 0;

QY 97 FGAASSYLNLEKLGEGSVATVYKGISRINGQVAVKVISMAEBGVFTAIRBASILKGL 156
Db 1 FGAADSYELKELGEGSVATVYKGSKVNGKLVALKVRLOEBEGTPTAIRBASILKGL 60
QY 157 KHANIVLHDIHTKETLTFVEFYMATDLAQYMSQHFGGLHPHNVRLFMFOLLRLGIATYH 216
Db 61 KHANIVLHDIHTKETLTFVEFYVHTDLQYMDKHPGGLHPDNYKLFLFOLLRLGISYH 120
QY 217 HOHVLHRLDKPONLISHLGEIKLADFGARAKSIPSTYSSSEVVTLMYRPPDALLGATE 276
Db 121 QRYILHRLDKPONLISHLGEIKLADFGARAKSVPSHTYSNEVVTLMYRPPDALLGATE 180
QY 277 YSSELDIWAGCIFIEMFGQPLFPQVSNILBQLEKIMEVLGVPTEDTPGVSKLPYNP 336
Db 181 YSTCLDMGVGCIFFVEMIGVAAFPQMKDIQDLERIFVLGTPNEDTWPVHSLPHFKP 240
QY 337 EWFPLPTPSLSLVNWRRLGRVPEADLASQMLKGFPRDVSAQELVHDYFSLPSQLYQ 396
Db 241 ERFLLSSSKNLQANWKLSYVNAEDLAKLQCSFKNRLSAQALSHYFSDLPRLWE 300
QY 397 LPDESLFTVSGVRLKPEKCDLLASYQKH 426
Db 301 LTDMSISFTVPVNRLOPEAGESMRAFGKN 330

Search completed: December 27, 2004, 12:25:43
Job time : 356 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2004, 11:08:34 ; Search time 24 Seconds
(without alignments)
1202.014 Million cell updates/sec

Title: US-10-786-065-5
Perfect score: 2330
Sequence: 1 MGDELCAKTVPQGCSCYHCS.....CDLASVQKGMHPAPFSKCM 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2330	100.0	435	4	US-09-801-861-5 Sequence 5, App1
2	2330	100.0	435	4	US-10-224-562-5 Sequence 5, App1
3	1784	76.9	343	4	US-09-801-861-2 Sequence 2, App1
4	1784	76.9	343	4	US-10-224-562-2 Sequence 2, App1
5	1195	51.5	330	4	US-09-801-861-8 Sequence 8, App1
6	1195	51.5	330	4	US-10-224-562-8 Sequence 8, App1
7	1194.5	51.5	423	4	US-09-206-344-2 Sequence 2, App1
8	1191.5	51.3	423	4	US-09-206-344-4 Sequence 4, App1
9	1191	51.3	330	4	US-09-801-861-9 Sequence 9, App1
10	1191	51.3	330	4	US-10-224-562-9 Sequence 9, App1
11	1190	51.3	330	4	US-09-801-861-10 Sequence 10, App1
12	1190	51.3	330	4	US-10-224-562-10 Sequence 10, App1
13	1000	43.1	245	4	US-09-801-861-7 Sequence 7, App1
14	1000	43.1	245	4	US-10-224-562-7 Sequence 7, App1
15	993	42.8	240	4	US-09-801-861-6 Sequence 6, App1
16	993	42.8	240	4	US-10-224-562-6 Sequence 6, App1
17	916	39.5	496	4	US-09-538-092-1239 Sequence 1239, App1
18	773.5	33.3	332	4	US-09-248-796A-18399 Sequence 18399, App1
19	751	32.4	305	4	US-09-538-092-771 Sequence 771, App1
20	743	32.0	292	1	US-08-154-915-2 Sequence 2, App1
21	743	32.0	292	2	US-08-154-915-3 Sequence 3, App1
22	743	32.0	292	2	US-08-246-361A-38 Sequence 38, App1
23	743	32.0	292	2	US-08-463-772-38 Sequence 38, App1
24	743	32.0	292	4	US-09-538-092-1238 Sequence 1238, App1
25	743	32.0	292	5	PCT-US93-09945-2 Sequence 2, App1
26	734	31.6	305	4	US-09-538-092-1236 Sequence 1236, App1
27	691.5	29.8	294	2	US-08-874-347-26 Sequence 26, App1

28	691.5	29.8	294	3	US-09-093-523-26 Sequence 26, App1
29	685	29.5	544	4	US-09-417-197-115 Sequence 115, App1
30	681.5	29.4	544	4	US-09-417-197-113 Sequence 113, App1
31	680	29.3	298	2	US-08-874-347-25 Sequence 25, App1
32	680	29.3	298	2	US-08-969-106-2 Sequence 2, App1
33	680	29.3	298	3	US-09-093-522-25 Sequence 25, App1
34	680	29.3	298	4	US-09-338-125-2 Sequence 2, App1
35	680	29.3	298	4	US-09-266-225D-14 Sequence 14, App1
36	679	29.3	298	3	US-09-457-040B-29 Sequence 29, App1
37	679	29.3	298	4	US-09-538-092-1006 Sequence 1006, App1
38	676	29.1	298	4	US-09-411-628-13 Sequence 13, App1
39	676	29.1	298	4	US-10-174-794-13 Sequence 13, App1
40	662.5	28.6	317	1	US-08-463-090B-9 Sequence 9, App1
41	662.5	28.6	317	2	US-08-874-347-18 Sequence 18, App1
42	662.5	28.6	317	3	US-09-093-522-18 Sequence 18, App1
43	662.5	28.6	317	3	US-09-248-796A-18426 Sequence 18426, App1
44	657.5	28.3	300	2	US-08-874-347-10 Sequence 10, App1
45	657.5	28.3	300	3	US-09-093-522-10 Sequence 10, App1

ALIGNMENTS

```
RESULT 1
US-09-801-861-5
; Sequence 5, Application US/09801861
; Patent No. 6492154
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01098
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Human
; US-09-801-861-5

Query Match      100.0%; Score 2320; DB 4; length 435;
Best Local Similarity 100.0%; Pred. No. 3.6e-226;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDELCAKTVPQGCSCYHCSGEGEHAHSCRRSQPETTEAFAFKLTDLKEASCMTSPHPRGL 60
DB 1 MGDELCAKTVPQGCSCYHCSGEGEHAHSCRRSQPETTEAFAFKLTDLKEASCMTSPHPRGL 60
QY 61 QAAKQKFKSKRRNSDCFOEEDLROGFOWKKSIPFGAASSYLNLKLGESGYATVYKG 120
DB 61 QAAKQKFKSKRRNSDCFOEEDLROGFOWKKSIPFGAASSYLNLKLGESGYATVYKG 120
QY 121 IIRINGQVAVLVKVISMAEEGVPTAIRBASILKGIKXANIVLHDITHTKETTFTVEY 180
DB 121 IIRINGQVAVLVKVISMAEEGVPTAIRBASILKGIKXANIVLHDITHTKETTFTVEY 180
QY 181 MTITDAQVNSQHPGGLHPNNVLFEMFOLLRGALVYIHQVLRHDKPOULLST 180
DB 181 MTITDAQVNSQHPGGLHPNNVLFEMFOLLRGALVYIHQVLRHDKPOULLST 180
QY 241 ADFGLARAKSISQTYSSSEVTLWTRPPDALLGATESSSELDIMWAGCT 241
DB 241 ADFGLARAKSISQTYSSSEVTLWTRPPDALLGATESSSELDIMWAGCT 241
QY 301 PGVSNILEQLEKIWEVLGVPTEDTWPGVSKLPNNVPEWPLFTPR 301
DB 301 PGVSNILEQLEKIWEVLGVPTEDTWPGVSKLPNNVPEWPLFTPR 301
QY 361 EDLASOMKGFPRDRVSAOEALVHDFSLPQOLVQLPDEP 361
DB 361 EDLASOMKGFPRDRVSAOEALVHDFSLPQOLVQLPDEP 361
```

Db 361 EDLASQMLKGFPPDRVSQAQALVHDYFSALPQSLYQLPDEBSLFTVSGVRLKPEMCDLLA 420
QY 421 SYQKGHHPPAQSFCW 435
Db 421 SYQKGHHPPAQSFCW 435

RESULT 2

US-10-224-562-5
; Sequence 5, Application us/10224562
; Patent No. 6730506
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001098DIV
; CURRENT APPLICATION NUMBER: US/10/224,562
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-224-562-5

Query Match 100.0%; Score 2320; DB 4; Length 435;
Best Local Similarity 100.0%; Pred. No. 3,6e-226;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGQELCAKTVQPGSCCYHSGEGEAHSCRRSQPETTEAAFKLTDLKEASCSMTSFHPRL 60
Db 1 MGQELCAKTVQPGSCCYHSGEGEAHSCRRSQPETTEAAFKLTDLKEASCSMTSFHPRL 60
QY 61 QAARAQKFKSKPRNSNDQFQEDLRQGFQWRKSLPFGAASSYLNLKLGEGSYATVYKG 120
Db 61 QAARAQKFKSKPRNSNDQFQEDLRQGFQWRKSLPFGAASSYLNLKLGEGSYATVYKG 120
QY 121 ISRINGQVALVKVISMNAEAGVPTAIRASLLKGLKHANIVLHDIHTKETLTFVPEY 180
Db 121 ISRINGQVALVKVISMNAEAGVPTAIRASLLKGLKHANIVLHDIHTKETLTFVPEY 180
QY 181 MHTDLAQYMSQHPGGLHPHNVRLFMFQLLRGLAYIHQHVLRDLKPQNLISHLGEELK 240
Db 181 MHTDLAQYMSQHPGGLHPHNVRLFMFQLLRGLAYIHQHVLRDLKPQNLISHLGEELK 240
QY 241 ADFGLARAKSIPQYSSSVTLWYRPPDALLGATEYSELDIWAGCIFIEMFGQPLF 300
Db 241 ADFGLARAKSIPQYSSSVTLWYRPPDALLGATEYSELDIWAGCIFIEMFGQPLF 300
QY 301 PGVSNILEQLEKIMEVLGVPTEDTWPVSKLPNYNDEWFPPLPTPSLHVVMNLGVPEA 360
Db 301 PGVSNILEQLEKIMEVLGVPTEDTWPVSKLPNYNDEWFPPLPTPSLHVVMNLGVPEA 360
QY 361 EDLASQMLKGFPPDRVSQAQALVHDYFSALPQSLYQLPDEBSLFTVSGVRLKPEMCDLLA 420
Db 361 EDLASQMLKGFPPDRVSQAQALVHDYFSALPQSLYQLPDEBSLFTVSGVRLKPEMCDLLA 420
QY 421 SYQKGHHPPAQSFCW 435
Db 421 SYQKGHHPPAQSFCW 435

RESULT 3

US-09-801-861-2
; Sequence 2, Application us/09801861
; Patent No. 6432154
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Human
US-09-801-861-2

Query Match 76.9%; Score 1784; DB 4; Length 343;
Best Local Similarity 100.0%; Pred. No. 4,8e-172;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGQELCAKTVQPGSCCYHSGEGEAHSCRRSQPETTEAAFKLTDLKEASCSMTSFHPRL 60
Db 1 MGQELCAKTVQPGSCCYHSGEGEAHSCRRSQPETTEAAFKLTDLKEASCSMTSFHPRL 60
QY 61 QAARAQKFKSKPRNSNDQFQEDLRQGFQWRKSLPFGAASSYLNLKLGEGSYATVYKG 120
Db 61 QAARAQKFKSKPRNSNDQFQEDLRQGFQWRKSLPFGAASSYLNLKLGEGSYATVYKG 120
QY 121 ISRINGQVALVKVISMNAEAGVPTAIRASLLKGLKHANIVLHDIHTKETLTFVPEY 180
Db 121 ISRINGQVALVKVISMNAEAGVPTAIRASLLKGLKHANIVLHDIHTKETLTFVPEY 180
QY 181 MHTDLAQYMSQHPGGLHPHNVRLFMFQLLRGLAYIHQHVLRDLKPQNLISHLGEELK 240
Db 181 MHTDLAQYMSQHPGGLHPHNVRLFMFQLLRGLAYIHQHVLRDLKPQNLISHLGEELK 240
QY 241 ADFGLARAKSIPQYSSSVTLWYRPPDALLGATEYSELDIWAGCIFIEMFGQPLF 300
Db 241 ADFGLARAKSIPQYSSSVTLWYRPPDALLGATEYSELDIWAGCIFIEMFGQPLF 300
QY 301 PGVSNILEQLEKIMEVLGVPTEDTWPVSKLPNYNP 336
Db 301 PGVSNILEQLEKIMEVLGVPTEDTWPVSKLPNYNP 336

RESULT 4

US-10-224-562-2
; Sequence 2, Application us/10224562
; Patent No. 6730506
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001098DIV
; CURRENT APPLICATION NUMBER: US/10/224,562
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-224-562-2

Query Match 76.9%; Score 1784; DB 4; Length 343;
Best Local Similarity 100.0%; Pred. No. 4,8e-172;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGQELCAKTVQPGSCCYHSGEGEAHSCRRSQPETTEAAFKLTDLKEASCSMTSFHPRL 60
Db 1 MGQELCAKTVQPGSCCYHSGEGEAHSCRRSQPETTEAAFKLTDLKEASCSMTSFHPRL 60
QY 61 QAARAQKFKSKPRNSNDQFQEDLRQGFQWRKSLPFGAASSYLNLKLGEGSYATVYKG 120
Db 61 QAARAQKFKSKPRNSNDQFQEDLRQGFQWRKSLPFGAASSYLNLKLGEGSYATVYKG 120
QY 121 ISRINGQVALVKVISMNAEAGVPTAIRASLLKGLKHANIVLHDIHTKETLTFVPEY 180

```
Db 121 ISINQVALAKVISMAEGVFTAIRBSLKLKLANIVLHDIHTKETLTVEFY 180
Qy 181 MHTDLAQYMSQHFGGLHPHNVRLFMQLRGLAYIHQHVHLRDLKPNLLISHLGELK 240
Db 181 MHTDLAQYMSQHFGGLHPHNVRLFMQLRGLAYIHQHVHLRDLKPNLLISHLGELK 240
Qy 241 ADPGLARASISPSQTSSEVYTLWYRPPDALLGATYSSELDITWAGCIFIEMFGQPLF 300
Db 241 ADPGLARASISPSQTSSEVYTLWYRPPDALLGATYSSELDITWAGCIFIEMFGQPLF 300
Qy 301 PGVSNILEQLEKIMEVLGVPTEDTWPVSKLPYNP 336
Db 301 PGVSNILEQLEKIMEVLGVPTEDTWPVSKLPYNP 336

RESULT 5
US-09-801-861-8
; Sequence 8, Application US/09801861
; Patent No. 6432154
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-801-861-8

Query Match 51.5%; Score 1195; DB 4; Length 330;
Best Local Similarity 68.2%; Pred. No. 2e-112;
Matches 225; Conservative 45; Mismatches 60; Indels 0; Gaps 0;

Qy 97 FGAASSYLNLEKLGSGSYATVYKGISRINGQVALKVISMAEGVFTAIRBSLKLK 156
Db 1 FGAADSYEKLKLGSGSYATVYKGSKNVGLKVALKVIHQEBEGTPTAIRBSLKLK 60
Qy 157 KLANIVLHDIHTKETLTVEFYMHTDLAQYMSQHFGGLHPHNVRLFMQLRGLAYIH 216
Db 61 KLANIVLHDIHTKETLTVEFYMHTDLAQYMSQHFGGLHPHNVRLFMQLRGLAYIH 120
Qy 217 HQHVHLRDLKPNLLISHLGELKADFGIARAKSIPSQTSSEVYTLWYRPPDALLGATE 276
Db 121 QRYIHLRDLKPNLLISHLGELKADFGIARAKSVSHSYNEVYTLWYRPPDALLGATE 180
Qy 277 YSELDITWAGCIFIEMFGQPLFPGVSNILEQLEKIMEVLGVPTEDTWPVSKLPYNP 336
Db 181 YSTCDLMQVGCIFEMFGQVAFPMKDIQDLERIFVLGTPEMDTWPVSHSLPHFNP 240
Qy 337 EWFPLPTPSLHVNNRLGRVBEADLASQMLKGFPRDRVSAQELVHDFGSLPSQLYQ 396
Db 241 ERFYVSSKSLQAWNKLSYVNHAEADLASKLQCSFKRSLSAQALSHFYFSDLPRLWE 300
Qy 397 LPDESILPTVSGVRLKPEMCDLLASYQKH 426
Db 301 LTDMSSIFTVPNVRLQPEAGSMRAFQKN 330

RESULT 6
US-10-224-562-8
; Sequence 8, Application US/10224562
; Patent No. 6730506
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
```

```
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001098DIY
; CURRENT APPLICATION NUMBER: US/10/224,562
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-224-562-8

Query Match 51.5%; Score 1195; DB 4; Length 330;
Best Local Similarity 68.2%; Pred. No. 2e-112;
Matches 225; Conservative 45; Mismatches 60; Indels 0; Gaps 0;

Qy 97 FGAASSYLNLEKLGSGSYATVYKGISRINGQVALKVISMAEGVFTAIRBSLKLK 156
Db 1 FGAADSYEKLKLGSGSYATVYKGSKNVGLKVALKVIHQEBEGTPTAIRBSLKLK 60
Qy 157 KLANIVLHDIHTKETLTVEFYMHTDLAQYMSQHFGGLHPHNVRLFMQLRGLAYIH 216
Db 61 KLANIVLHDIHTKETLTVEFYMHTDLAQYMSQHFGGLHPHNVRLFMQLRGLAYIH 120
Qy 217 HQHVHLRDLKPNLLISHLGELKADFGIARAKSIPSQTSSEVYTLWYRPPDALLGATE 276
Db 121 QRYIHLRDLKPNLLISHLGELKADFGIARAKSVSHSYNEVYTLWYRPPDALLGATE 180
Qy 277 YSELDITWAGCIFIEMFGQPLFPGVSNILEQLEKIMEVLGVPTEDTWPVSKLPYNP 336
Db 181 YSTCDLMQVGCIFEMFGQVAFPMKDIQDLERIFVLGTPEMDTWPVSHSLPHFNP 240
Qy 337 EWFPLPTPSLHVNNRLGRVBEADLASQMLKGFPRDRVSAQELVHDFGSLPSQLYQ 396
Db 241 ERFYVSSKSLQAWNKLSYVNHAEADLASKLQCSFKRSLSAQALSHFYFSDLPRLWE 300
Qy 397 LPDESILPTVSGVRLKPEMCDLLASYQKH 426
Db 301 LTDMSSIFTVPNVRLQPEAGSMRAFQKN 330

RESULT 7
US-09-206-344A-2
; Sequence 2, Application US/09206344A
; Patent No. 6432668
; GENERAL INFORMATION:
; APPLICANT: Christoph Reinhard
; APPLICANT: David Pot
; APPLICANT: Altaf Kassem
; APPLICANT: Tasha Marenbach
; APPLICANT: Lewis T. Williams
; TITLE OF INVENTION: HUMAN CYCLIN-DEPENDENT KINASE (HPPFAIRE)
; FILE REFERENCE: 200103.445 / 1429.002
; CURRENT APPLICATION NUMBER: US/09/206,344A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-206-344A-2

Query Match 51.5%; Score 1194.5; DB 4; Length 423;
Best Local Similarity 60.5%; Pred. No. 3.3e-112;
Matches 234; Conservative 55; Mismatches 83; Indels 15; Gaps 2;

Qy 51 SMTSFHPRGLQAAARQKSKRPRSNSDCFOEEDLRQGFQWRK-----SLPFGA 99
Db 30 TOSTFDPKPKANQKRVHSE-----NNACINFKTSSTGKSPKVRHSSPSPSPKFGK 85
Qy 100 ASSYIANLEKLGSGSYATVYKGISRINGQVALKVISMAEGVFTAIRBSLKLKGLGA 159
```

```
Db      86 ADSYEKLEKLGSGSYATVYKGSRVNGKLVALKVIRLOEBEGTPTTAIRASILKGLKHA 145
Qy      160 NIVLHDIHTKXETLTFVEEYMTDLAQYMSQHPGGLHPHNVRLEFMPOLLRLGLAYIHHOH 219
      146 NIVLHDIHTKXETLTFVEEYVHTDLCQYMDKHFGSLHNDVNYKLFOLLRLGLSYIHHQY 205
Qy      220 VLRDLKPNOLLISHLGEKLADFGILARAKSIPSQTSSEVYTLWYRPPDALLGATEYSS 279
      206 ILRDLKPNOLLISDTGELKADFGILARAKSVSHTSYNSNEVYTLWYRPPDVLGSTEYST 265
Qy      280 ELDINGAGCIFIEMFGOGPLFPGVSNILEOLEKIMVLCVPTEDTPGVSXKLPNNPEMF 339
      266 CLDMGVGCI FVEEMIQGVAAPFGMKDIOQLERIFLVLTGTPMEDTPGVHSLPHFKPERF 325
Qy      340 PLPTPSLHVNNRLGRVPEAEADLASQMLKGFPRDRVSAQELVHDYFSALEPSQLYQLPD 399
      326 TLVSSKNLQAVNNKLSYVNHAEADLASKLLQCSFKNRLSQAALLSHYFSDLPRLMELTD 385
Qy      400 EESLFTVSGVRLKPEMCDLLASYOKGH 426
      386 MSSIFTVPNVRLQPEAGESMRAFGKN 412
```

```
RESULT 8
US-09-206-344A-4
/ Sequence 4, Application US/09206344A
/ Patent No. 6432668
/ GENERAL INFORMATION:
/ APPLICANT: Christoph Reinhard
/ APPLICANT: David Pot
/ APPLICANT: Alcat Kassar
/ APPLICANT: Tasha Marenbach
/ APPLICANT: Lewis T. Williams
/ TITLE OF INVENTION: HUMAN CYCLIN-DEPENDENT KINASE (HPTAIRE)
/ FILE REFERENCE: 200103.445 / 1429.002
/ CURRENT APPLICATION NUMBER: US/09/206,344A
/ CURRENT FILING DATE: 1998-12-17
/ NUMBER OF SEQ ID NOS: 13
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 4
/ LENGTH: 423
/ TYPE: PRT
/ ORGANISM: mouse
US-09-206-344A-4
```

```
Query Match      51.4%; Score 1191.5; DB 4; Length 423;
Best Local Similarity 60.2%; Pred. No. 6,6e-112;
Matches 233; Conservative 55; Mismatches 84; Indels 15; Gaps 2;
```

```
Qy      51 SMTSFHPRGLQAAKAKFKSKRPRNSDCFOEBDLRQGFQWRK-----SLPFGA 99
      30 TGSTPDPPEKPAQVQRVHSE---NNACINFKSSAGKESPKVRBHSPTSPTSPKFGK 85
Qy      100 ASSYMLLEKLGSGSYATVYKGSRVNGKLVALKVIRLOEBEGTPTTAIRASILKGLKHA 159
      86 ADSYEKLEKLGSGSYATVYKGSRVNGKLVALKVIRLOEBEGTPTTAIRASILKGLKHA 145
Qy      160 NIVLHDIHTKXETLTFVEEYMTDLAQYMSQHPGGLHPHNVRLEFMPOLLRLGLAYIHHOH 219
      146 NIVLHDIHTKXETLTFVEEYVHTDLCQYMDKHFGSLHNDVNYKLFOLLRLGLSYIHHQY 205
Qy      220 VLRDLKPNOLLISHLGEKLADFGILARAKSIPSQTSSEVYTLWYRPPDALLGATEYSS 279
      206 ILRDLKPNOLLISDTGELKADFGILARAKSVSHTSYNSNEVYTLWYRPPDVLGSTEYST 265
Qy      280 ELDINGAGCIFIEMFGOGPLFPGVSNILEOLEKIMVLCVPTEDTPGVSXKLPNNPEMF 339
      266 CLDMGVGCI FVEEMIQGVAAPFGMKDIOQLERIFLVLTGTPMEDTPGVHSLPHFKPERF 325
Qy      340 PLPTPSLHVNNRLGRVPEAEADLASQMLKGFPRDRVSAQELVHDYFSALEPSQLYQLPD 399
      326 TLVSSKNLQAVNNKLSYVNHAEADLASKLLQCSFKNRLSQAALLSHYFSDLPRLMELTD 385
```

```
Qy      400 EESLFTVSGVRLKPEMCDLLASYOKGH 426
      386 MSSIFTVPNVRLQPEAGESMRAFGKN 412
```

```
RESULT 9
US-09-801-861-9
/ Sequence 9, Application US/09801861
/ Patent No. 6492154
/ GENERAL INFORMATION:
/ APPLICANT: YAN, Chunhua et al.
/ TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
/ TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
/ TITLE OF INVENTION: THEREOF
/ FILE REFERENCE: CL001098
/ CURRENT APPLICATION NUMBER: US/09/801,861
/ CURRENT FILING DATE: 2001-03-09
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 9
/ LENGTH: 330
/ TYPE: PRT
/ ORGANISM: Human
US-09-801-861-9
```

```
Query Match      51.3%; Score 1191; DB 4; Length 330;
Best Local Similarity 67.9%; Pred. No. 5e-112;
Matches 224; Conservative 46; Mismatches 60; Indels 0; Gaps 0;
```

```
Qy      97 FGAASSYMLLEKLGSGSYATVYKGSRVNGKLVALKVIRLOEBEGTPTTAIRASILKGL 156
      1 FKAADSYKLEKLGSGSYATVYKGSRVNGKLVALKVIRLOEBEGTPTTAIRASILKGL 60
Qy      157 KHANIVLHDIHTKXETLTFVEEYMTDLAQYMSQHPGGLHPHNVRLEFMPOLLRLGLAYI 216
      61 KHANIVLHDIHTKXETLTFVEEYVHTDLCQYMDKHFGSLHNDVNYKLFOLLRLGLSYI 120
Qy      217 HOHVLHRLDKPNOLLISHLGEKLADFGILARAKSIPSQTSSEVYTLWYRPPDALLGATE 276
      121 QRYIIRDLKPNOLLISDTGELKADFGILARAKSVSHTSYNSNEVYTLWYRPPDVLGSTE 180
Qy      227 YSELDINGAGCIFIEMFGOGPLFPGVSNILEOLEKIMVLCVPTEDTPGVSXKLPNNP 336
      181 YSTCLDMGVGCI FVEEMIQGVAAPFGMKDIOQLERIFLVLTGTPMEDTPGVHSLPHFKP 240
Qy      337 EWFPLPTPSLHVNNRLGRVPEAEADLASQMLKGFPRDRVSAQELVHDYFSALEPSQLY 396
      241 ERFITYSSKNLQAVNNKLSYVNHAEADLASKLLQCSFKNRLSQAALLSHYFSDLPRLME 300
Qy      397 LPDESLFTVSGVRLKPEMCDLLASYOKGH 426
      301 LTDMSSIFTVPNVRLQPEAGESMRAFGKN 330
```

```
RESULT 10
US-10-224-562-9
/ Sequence 9, Application US/10224562
/ Patent No. 6730506
/ GENERAL INFORMATION:
/ APPLICANT: YAN, Chunhua et al.
/ TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
/ TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
/ TITLE OF INVENTION: THEREOF
/ FILE REFERENCE: CL001098DIV
/ CURRENT APPLICATION NUMBER: US/10/224,562
/ CURRENT FILING DATE: 2002-08-21
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 9
/ LENGTH: 330
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-224-562-9
```

```

Query Match      51.3%; Score 1190; DB 4; Length 330;
Best Local Similarity 67.9%; Pred. No. 5e-112;
Matches 224; Conservative 46; Mismatches 60; Indels 0; Gaps 0;

QY 97 FGAASSYLNLEKLGSGSYATVYKGSIRINGOLVALKVISMNAEEGVPTAIRBASLLKGL 156
DB 1 FGAADSYELKELGSGSYATVYKGSKVNGKLVALKVIRLOEEBGPPTAIRBASLLKGL 60
QY 157 KIANIVLHDIHTKETTLPFEVYVMTDLAQYMSQHPGLHPHNVLFWFOLLRLGLAYIH 216
DB 61 KIANIVLHDIHTKETTLPFEVYVMTDLAQYMDKPGGHPNDVNLFLFOLLRLGLSYIH 120
QY 217 HQVILHRDLKPNLLISHLGEIKLADFGIARAKSIPSQTSSEVVTLWYRPPDALIGATE 276
DB 121 QRYIILHRDLKPNLLISDTGELKLDGFLARAKSVSHSYNSNEVTLWYRPPDVLLGSTE 180
QY 277 YSESLDIMGAGCIFIEMFGQPLFPGVSNILOLEKIMEVLGVPTEDTPGVSKLPNNVP 336
DB 181 YSTCLDMGVGCI FVEIMIGVAAFPGMKIIOQLERIFLVLTGPNEDTPGVHSLPHFKP 240
QY 337 EWFPLPTPSLSLVNVRRLGRVPEABDLASQMLKGPRDRVSAQELVHDYFSLPSQLYQ 396
DB 241 ERFITYNSKSLRQANNKLSYVNHABDLASKLQCSFKRLSQAALSHYFSDLPRLWE 300
QY 397 LPDESLFTVSGVRLKPEWCDDLASQYOKGH 426
DB 301 LTDMSIFTVPNVRLQPEAGESMRAFGKNN 330

RESULT 11
US-09-801-861-10
; Sequence 10, Application US/09801861
; Patent No. 6492154
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-801-861-10

Query Match      51.3%; Score 1190; DB 4; Length 330;
Best Local Similarity 67.9%; Pred. No. 6.4e-112;
Matches 224; Conservative 45; Mismatches 61; Indels 0; Gaps 0;

QY 97 FGAASSYLNLEKLGSGSYATVYKGSIRINGOLVALKVISMNAEEGVPTAIRBASLLKGL 156
DB 1 FGAADSYELKELGSGSYATVYKGSKVNGKLVALKVIRLOEEBGPPTAIRBASLLKGL 60
QY 157 KIANIVLHDIHTKETTLPFEVYVMTDLAQYMSQHPGLHPHNVLFWFOLLRLGLAYIH 216
DB 61 KIANIVLHDIHTKETTLPFEVYVMTDLAQYMDKPGGHPNDVNLFLFOLLRLGLSYIH 120
QY 217 HQVILHRDLKPNLLISHLGEIKLADFGIARAKSIPSQTSSEVVTLWYRPPDALIGATE 276
DB 121 QRYIILHRDLKPNLLISDTGELKLDGFLARAKSVSHSYNSNEVTLWYRPPDVLLGSTE 180
QY 277 YSESLDIMGAGCIFIEMFGQPLFPGVSNILOLEKIMEVLGVPTEDTPGVSKLPNNVP 336
DB 181 YSTCLDMGVGCI FVEIMIGVAAFPGMKIIOQLERIFLVLTGPNEDTPGVHSLPHFKP 240
QY 337 EWFPLPTPSLSLVNVRRLGRVPEABDLASQMLKGPRDRVSAQELVHDYFSLPSQLYQ 396
DB 241 ERFITYNSKSLRQANNKLSYVNHABDLASKLQCSFKRLSQAALSHYFSDLPRLWE 300

```

```

QY 397 LPDESLFTVSGVRLKPEWCDDLASQYOKGH 426
DB 301 LTDMSIFTVPNVRLQPEAGESMRAFGKNN 330

RESULT 12
US-10-224-562-10
; Sequence 10, Application US/10224562
; Patent No. 6730506
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01098DIV
; CURRENT APPLICATION NUMBER: US/10/224,562
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-224-562-10

Query Match      51.3%; Score 1190; DB 4; Length 330;
Best Local Similarity 67.9%; Pred. No. 6.4e-112;
Matches 224; Conservative 45; Mismatches 61; Indels 0; Gaps 0;

QY 97 FGAASSYLNLEKLGSGSYATVYKGSIRINGOLVALKVISMNAEEGVPTAIRBASLLKGL 156
DB 1 FGAADSYELKELGSGSYATVYKGSKVNGKLVALKVIRLOEEBGPPTAIRBASLLKGL 60
QY 157 KIANIVLHDIHTKETTLPFEVYVMTDLAQYMSQHPGLHPHNVLFWFOLLRLGLAYIH 216
DB 61 KIANIVLHDIHTKETTLPFEVYVMTDLAQYMDGHPGLHPNDVNLFLFOLLRLGLSYIH 120
QY 217 HQVILHRDLKPNLLISHLGEIKLADFGIARAKSIPSQTSSEVVTLWYRPPDALIGATE 276
DB 121 QRYIILHRDLKPNLLISDTGELKLDGFLARAKSVSHSYNSNEVTLWYRPPDVLLGSTE 180
QY 277 YSESLDIMGAGCIFIEMFGQPLFPGVSNILOLEKIMEVLGVPTEDTPGVSKLPNNVP 336
DB 181 YSTCLDMGVGCI FVEIMIGVAAFPGMKIIOQLERIFLVLTGPNEDTPGVHSLPHFKP 240
QY 337 EWFPLPTPSLSLVNVRRLGRVPEABDLASQMLKGPRDRVSAQELVHDYFSLPSQLYQ 396
DB 241 ERFITYNSKSLRQANNKLSYVNHABDLASKLQCSFKRLSQAALSHYFSDLPRLWE 300

QY 397 LPDESLFTVSGVRLKPEWCDDLASQYOKGH 426
DB 301 LTDMSIFTVPNVRLQPEAGESMRAFGKNN 330

RESULT 13
US-09-801-861-7
; Sequence 7, Application US/09801861
; Patent No. 6492154
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Human

```

US-09-801-861-7

Query Match 43.1%; Score 1000; DB 4; Length 245;

Best Local Similarity 75.9%; Pred. No. 7e-93;

Matches 186; Conservative 25; Mismatches 34; Indels 0; Gaps 0;

QY 97 FGAASSYLNLEKLGEGSYATVYKGISRINGQVAVLKVISMAEAGVPFTAIRBASLLKGL 156
 DB 1 FGAADSYELKELGEGSYATVYKGSXNGKLVALKVIRLQEBEGTPTAIRBASLLKGL 60
 QY 157 KXANIVLHLDIHTKETTTFVEFYMHDTDLAQYMSQHPGGLHPHNVRLFMFOLLGLAYIH 216
 DB 61 KXANIVLHLDIHTKETTTLVFEYVHTDLCQYMDKHPGGLHPNVKFLFOLLRLGLSYIH 120
 QY 217 HOHVLRDLKPNOLLISHLGEIKLADFGIARAKSIPSOYSSSEVVTLMWRPDPALLGATE 276
 DB 121 QRYIILHRDLKPNOLLISDTGELKADFGIARAKSVPSHTYSNEVVTLMWRPDPVLLGSTE 180
 QY 277 YSESLDIWAGGCIETEMFGQGPLFPGVSNILBQLEKIMEVLGVPTEDTWPVSKLPNNYP 336
 DB 181 YSTCLDMWGVGCIETEMIGVAAFPQMKDIQDLERIFVLGTPTNEDTWPVHSLPHFKP 240
 QY 337 EWFPL 341
 DB 241 ERFTL 245

RESULT 14

US-10-224-562-7

Sequence 7; Application US/10224562

Patent No. 6730506

GENERAL INFORMATION:

APPLICANT: YAN, Chunhua et al.

TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

FILE REFERENCE: CL001098DIV

CURRENT APPLICATION NUMBER: US/10/224,562

CURRENT FILING DATE: 2002-08-21

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 245

TYPE: PRT

ORGANISM: Homo sapiens

US-10-224-562-7

Query Match 43.1%; Score 1000; DB 4; Length 245;

Best Local Similarity 75.9%; Pred. No. 7e-93;

Matches 186; Conservative 25; Mismatches 34; Indels 0; Gaps 0;

QY 97 FGAASSYLNLEKLGEGSYATVYKGISRINGQVAVLKVISMAEAGVPFTAIRBASLLKGL 156
 DB 1 FGAADSYELKELGEGSYATVYKGSXNGKLVALKVIRLQEBEGTPTAIRBASLLKGL 60
 QY 157 KXANIVLHLDIHTKETTTFVEFYMHDTDLAQYMSQHPGGLHPHNVRLFMFOLLGLAYIH 216
 DB 61 KXANIVLHLDIHTKETTTLVFEYVHTDLCQYMDKHPGGLHPNVKFLFOLLRLGLSYIH 120
 QY 217 HOHVLRDLKPNOLLISHLGEIKLADFGIARAKSIPSOYSSSEVVTLMWRPDPALLGATE 276
 DB 121 QRYIILHRDLKPNOLLISDTGELKADFGIARAKSVPSHTYSNEVVTLMWRPDPVLLGSTE 180
 QY 277 YSESLDIWAGGCIETEMFGQGPLFPGVSNILBQLEKIMEVLGVPTEDTWPVSKLPNNYP 336
 DB 181 YSTCLDMWGVGCIETEMIGVAAFPQMKDIQDLERIFVLGTPTNEDTWPVHSLPHFKP 240
 QY 337 EWFPL 341
 DB 241 ERFTL 245

RESULT 15

US-09-801-861-6

Sequence 6; Application US/09801861

Patent No. 6492154

GENERAL INFORMATION:

APPLICANT: YAN, Chunhua et al.

TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

FILE REFERENCE: CL001098

CURRENT APPLICATION NUMBER: US/09/801,861

CURRENT FILING DATE: 2001-03-09

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 240

TYPE: PRT

ORGANISM: Mus musculus

US-09-801-861-6

Query Match 42.8%; Score 993; DB 4; Length 240;

Best Local Similarity 76.7%; Pred. No. 3.5e-92;

Matches 184; Conservative 24; Mismatches 32; Indels 0; Gaps 0;

QY 97 FGAASSYLNLEKLGEGSYATVYKGISRINGQVAVLKVISMAEAGVPFTAIRBASLLKGL 156
 DB 1 FGAADSYELKELGEGSYATVYKGSXNGKLVALKVIRLQEBEGTPTAIRBASLLKGL 60
 QY 157 KXANIVLHLDIHTKETTTFVEFYMHDTDLAQYMSQHPGGLHPHNVRLFMFOLLGLAYIH 216
 DB 61 KXANIVLHLDIHTKETTTLVFEYVHTDLCQYMDKHPGGLHPNVKFLFOLLRLGLSYIH 120
 QY 217 HOHVLRDLKPNOLLISHLGEIKLADFGIARAKSIPSOYSSSEVVTLMWRPDPALLGATE 276
 DB 121 QRYIILHRDLKPNOLLISDTGELKADFGIARAKSVPSHTYSNEVVTLMWRPDPVLLGSTE 180
 QY 277 YSESLDIWAGGCIETEMFGQGPLFPGVSNILBQLEKIMEVLGVPTEDTWPVSKLPNNYP 336
 DB 181 YSTCLDMWGVGCIETEMIGVAAFPQMKDIQDLERIFVLGTPTNEDTWPVHSLPHFKP 240

Search completed: December 27, 2004, 12:19:38

Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2004, 04:18:24 ; Search time 78 Seconds
(without alignments)
3208.820 Million cell updates/sec

Title: US-10-786-065-5
Perfect score: 2320
Sequence: 1 MGDELCACTVPGGSCSYHCS.....CDLASYQKGMHPAPQSKCW 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues
Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2038	87.8	384	1 A2S7 HUMAN	Q96440 homo sapien
2	1844	79.5	349	2 Q81UP1	Q81UP1 homo sapien
3	1641.5	70.8	345	2 Q6ZMR9	Q6ZMR9 homo sapien
4	1641.5	70.8	345	2 BAD18656	BAD18656 homo sapi
5	1198.5	51.7	469	1 KPRT MOUSE	Q55495 mus musculu
6	1192.5	51.4	423	2 Q6NVF8	Q6NVF8 mus musculu
7	1192.5	51.4	423	2 AAH68134	AAH68134 mus muscu
8	1192.5	51.4	453	2 Q6ZQ37	Q6ZQ37 mus musculu
9	1192.5	51.4	453	2 BAC98036	BAC98036 mus muscu
10	1191	51.3	435	2 Q6DJM7	Q6DJM7 xenopus lae
11	1185.5	51.1	469	1 KPRT HUMAN	Q94921 homo sapien
12	1078.5	46.5	367	2 Q8WZ44	Q8WZ44 drosophila
13	1078.5	46.5	501	2 Q7KX04	Q7KX04 drosophila
14	1078.5	46.5	509	2 Q7KX08	Q7KX08 drosophila
15	1078.5	46.5	509	2 Q9VZNR	Q9VZNR drosophila
16	1078.5	46.5	522	2 Q7KX03	Q7KX03 drosophila
17	1078.5	46.5	522	2 Q9XK93	Q9XK93 drosophila
18	1074.5	46.3	509	2 Q94888	Q94888 drosophila
19	1067.5	46.0	446	2 Q7Q6Q3	Q7Q6Q3 anopheles
20	1024.5	44.2	339	2 Q8BN81	Q8BN81 mus musculu
21	1022.5	44.1	339	1 Q8C411	Q8C411 mus musculu
22	916	39.5	496	1 KPRT HUMAN	Q00536 homo sapien
23	916	39.5	496	1 AAP35473	AAP35473 homo sapi
24	912	39.3	496	1 KPRT MOUSE	Q04735 mus musculu
25	912	39.3	496	2 Q91WF4	Q91WF4 mus musculu
26	912	39.3	496	2 BAC34635	BAC34635 mus muscu
27	909.5	39.2	395	2 Q96GAS	Q96GAS homo sapien
28	909.5	39.2	448	2 Q9BRL4	Q9BRL4 homo sapien
29	908	39.1	496	1 KPRT RAT	Q61866 rattus norv
30	900	38.8	451	1 KPRT MOUSE	Q04899 mus musculu
31	900	38.8	451	2 BAB33732	BAB33732 mus muscu

32	897.5	38.7	504	2 Q6V3A2	Q6V3A2 homo sapien
33	897.5	38.7	504	2 AAR13066	AAR13066 homo sapi
34	897	38.7	523	2 Q8NEB8	Q8NEB8 homo sapien
35	896	38.6	451	1 KPRT RAT	Q35832 rattus norv
36	895	38.6	474	2 Q6V3A3	Q6V3A3 homo sapien
37	895	38.6	474	2 AAR13065	AAR13065 homo sapi
38	893	38.5	472	1 KPRT HUMAN	Q07002 homo sapien
39	893	38.5	472	2 AAP35963	AAP35963 homo sapi
40	893	38.5	523	1 KPRT MOUSE	Q8K0D0 mus musculu
41	891	38.4	500	2 Q7ZVJ1	Q7ZVJ1 xenopus lae
42	890	38.4	523	1 KPRT HUMAN	Q00537 homo sapien
43	885.5	38.2	523	1 KPRT RAT	Q35831 rattus norv
44	876	37.8	462	2 Q6DFZ8	Q6DFZ8 xenopus tro
45	821	35.4	551	2 Q45022	Q45022 caenorhabdi

ALIGNMENTS

RESULT 1
ID A2S7 HUMAN STANDARD; PRT; 384 AA.
AC Q96Q40;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Serine/threonine-protein kinase AHS2CR7 (EC 2.7.1.37) (Amyotrophic lateral sclerosis 2 chromosomal region candidate gene protein 7).
GN Name=ALS2CR7;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21470351; PubMed=11586298; DOI=10.1038/ng1001-166;
RA Hadano S., Hand C.K., Oeuga H., Yanagisawa Y., Otsu A., Devon R.S., Miyamoto N., Shouguuchi-Miyata J., Okada Y., Singara R., Figlewicz D.A., Kwiatkowski T., Hosler B.A., Sagie T., Skaug J., Nasir J., Brown R.H. Jr., Scherer S.W., Rouleau G.A., Hayden M.R., Ikeda J.-E.;
RA "A gene encoding a novel GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2";
RT Nat. Genet. 29:166-173(2001).
RL -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC CDCX/CDKX subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL: AB053308; BAB69017.1; -
CC HSSP: P24941; ID18.
CC Genew: HGNC:14434; ALS2CR7.
DR InterPro: IPR011009; Kinase_Like.
DR InterPro: IPR00719; Prot_Kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_kin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR Prodom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
FT DOMAIN 52 336
FT NP_BIND 58 66 ATP (By similarity).

```

FT BINDING      81      81      ATP (by similarity) .
FT ACT_SITE    173      173      Proton acceptor (By similarity) .
SQ SEQUENCE     384 AA; 43574 MW; 6CAD08B6E0E43EB CRC64;
Query Match      87.8%; Score 2038; DB 1; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.3e-137;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 MTSFHRGLOAARAKQFKSKRPRNSDCFOEEDLRQGFQWRKSLPRGAASSYLNEKIGE 111
DB 1 MTSFHRGLOAARAKQFKSKRPRNSDCFOEEDLRQGFQWRKSLPRGAASSYLNEKIGE 60
QY 112 GSATATYKGISIRINGQVLAALKVISMNAEAGVPTAIRASLSLKGKHNIVLHDIHTK 171
DB 61 GSATATYKGISIRINGQVLAALKVISMNAEAGVPTAIRASLSLKGKHNIVLHDIHTK 120
QY 172 ETLTFVFEYMTDLAQYMSQHPGGLPHNVRLFMFQLRLGLAYIHQVHLRDLKQNL 231
DB 121 ETLTFVFEYMTDLAQYMSQHPGGLPHNVRLFMFQLRLGLAYIHQVHLRDLKQNL 180
QY 232 ISHLGELKLADEGLARAKSIPSGTYSSEVVTLMWRPPDALLGATEYSELDTWGACIF 291
DB 181 ISHLGELKLADEGLARAKSIPSGTYSSEVVTLMWRPPDALLGATEYSELDTWGACIF 240
QY 292 EMFOGQPLPFGVSNILDELEKIWEVLGVPTEDTWPGVSKLPVYNPEWFPPLTPRSLHVV 351
DB 241 EMFOGQPLPFGVSNILDELEKIWEVLGVPTEDTWPGVSKLPVYNPEWFPPLTPRSLHVV 300
QY 352 NRIGRVPEAEADLASQMLKGFPDRDVSQAEALVHDYFSALPSQLYQLPDESELTFTVSGVRL 411
DB 301 NRIGRVPEAEADLASQMLKGFPDRDVSQAEALVHDYFSALPSQLYQLPDESELTFTVSGVRL 360
QY 412 KPEMCDLLASYOKGHHPAQFSKCM 435
DB 361 KPEMCDLLASYOKGHHPAQFSKCM 384

RESULT 2
O81UP1 PRELIMINARY; PRT; 349 AA.
AC O81UP1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ALS3CR7 protein.
OS Homo sapiens (human) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Schat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.U., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huljk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman W., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL: BC038607; AF38807.1; -.
DR HSSP: P2941; ID18.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0004668; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR001199; Cyt-B5.
DR InterPro: IPR011009; Kinase_like.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00191; CYTOCHROME_B5_1; UNKNOWN_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 349 AA; 39597 MW; F884A39447AaB66 CRC64;

Query Match      79.5%; Score 1844; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 8.6e-124;
Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 MTSFHRGLOAARAKQFKSKRPRNSDCFOEEDLRQGFQWRKSLPRGAASSYLNEKIGE 111
DB 1 MTSFHRGLOAARAKQFKSKRPRNSDCFOEEDLRQGFQWRKSLPRGAASSYLNEKIGE 60
QY 112 GSATATYKGISIRINGQVLAALKVISMNAEAGVPTAIRASLSLKGKHNIVLHDIHTK 171
DB 61 GSATATYKGISIRINGQVLAALKVISMNAEAGVPTAIRASLSLKGKHNIVLHDIHTK 120
QY 172 ETLTFVFEYMTDLAQYMSQHPGGLPHNVRLFMFQLRLGLAYIHQVHLRDLKQNL 231
DB 121 ETLTFVFEYMTDLAQYMSQHPGGLPHNVRLFMFQLRLGLAYIHQVHLRDLKQNL 180
QY 232 ISHLGELKLADEGLARAKSIPSGTYSSEVVTLMWRPPDALLGATEYSELDTWGACIF 291
DB 181 ISHLGELKLADEGLARAKSIPSGTYSSEVVTLMWRPPDALLGATEYSELDTWGACIF 240
QY 292 EMFOGQPLPFGVSNILDELEKIWEVLGVPTEDTWPGVSKLPVYNPEWFPPLTPRSLHVV 351
DB 241 EMFOGQPLPFGVSNILDELEKIWEVLGVPTEDTWPGVSKLPVYNPEWFPPLTPRSLHVV 300
QY 352 NRIGRVPEAEADLASQMLKGFPDRDVSQAEALVHDYFSALPSQLYQLPDE 400
DB 301 NRIGRVPEAEADLASQMLKGFPDRDVSQAEALVHDYFSALPSQLYQLPDE 349

RESULT 3
O6ZMR9 PRELIMINARY; PRT; 345 AA.
AC O6ZMR9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ16732.
OS Homo sapiens (human) .
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RP Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,

```

RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Magatsuna M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
 RA Suzuki Y., Sugano S., Nagahori K., Maehuo Y., Nagai K., Isegaki T.,
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL: AK131512; BAD18656.1; -.
 DR GO: GO:0016301; F-kinase activity; IEA.
 DR InterPro: IPR011009; Kinase-like.
 DR InterPro: IPR000719; Prot_Kinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR008271; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00219; TYRc; 1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE: PS00011; PROTEIN KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
 KW SEQUENCE 345 AA; 38586 MW; A51640CA8C119F43 CRC64;
 SQ

Query Match 70.8%; Score 1641.5; DB 2; Length 345;
 Best Local Similarity 93.8%; Pred. No. 2.6e-109;
 Matches 315; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

QY 1 MGELCAKTVQPCSCYHCSEGEAHSRRSOPETTEAAFKLTDLEKASCSMTSPHPRGL 60
 DB 1 MGELCAKTVQPCSCYHCSEGEAHSRRSOPETTEAAFKLTDLEKASCSMTSPHPRGL 60
 QY 61 QAAARQKFKSKPRNSDCFOEEDLRQGFQWRKSLPFGAASSTYLNLEKGBGSYATVYVG 120
 DB 61 QAAARQKFKSKPRNSDCFOEEDLRQGFQWRKSLPFGAASSTYLNLEKGBGSYATVYVG 120
 QY 121 ISRINGQVALKAVISNNAEEGVPTAIRASLLKGLKHNIVLHDIHTKETLTFVPEY 180
 DB 121 ISRINGQVALKAVISNNAEEGVPTAIRASLLKGLKHNIVLHDIHTKETLTFVPEY 180
 QY 181 MHTDLAQYNSQHPGGLPHNVLRFMFQQLRGLAYIHQVHLRDLKPNQLLSHIGELKL 240
 DB 181 M-----LRFQQLRGLAYIHQVHLRDLKPNQLLSHIGELKL 219
 QY 241 ADFGLARAKSISQYSSSEVTLWYRPPDALLGATYSESLDIWGAGCIFIEMFOQPLF 300
 DB 241 ADFGLARAKSISQYSSSEVTLWYRPPDALLGATYSESLDIWGAGCIFIEMFOQPLF 279
 QY 301 PGVSNILEQLEKIMEVLGVPTEDTWPGVSKLPNYP 336
 DB 301 PGVSNILEQLEKIMEVLGVPTEDTWPGVSKLPNYP 315

RESULT 4
 BAD18656 PRELIMINARY; PRT; 345 AA.
 AC BAD18656;
 DT 12-MAY-2004 (TRENBLrel. 27, Created)
 DT 12-MAY-2004 (TRENBLrel. 27, Last sequence update)
 DE CDNA FLJ16732 fig. clone BNGH42005017, moderately similar to
 DE SERINE/THREONINE-PROTEIN KINASE PCFAIRE-1 (EC 2.7.1.-).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hito Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Magatsuna M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
 RA Suzuki Y., Sugano S., Nagahori K., Maehuo Y., Nagai K., Isegaki T.,
 RT "NEBO human cDNA sequencing project.";

RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK131512; BAD18656.1; -.
 KW Kinase.
 SQ SEQUENCE 345 AA; 38586 MW; A51640CA8C119F43 CRC64;
 SQ

Query Match 70.8%; Score 1641.5; DB 2; Length 345;
 Best Local Similarity 93.8%; Pred. No. 2.6e-109;
 Matches 315; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

QY 1 MGELCAKTVQPCSCYHCSEGEAHSRRSOPETTEAAFKLTDLEKASCSMTSPHPRGL 60
 DB 1 MGELCAKTVQPCSCYHCSEGEAHSRRSOPETTEAAFKLTDLEKASCSMTSPHPRGL 60
 QY 61 QAAARQKFKSKPRNSDCFOEEDLRQGFQWRKSLPFGAASSTYLNLEKGBGSYATVYVG 120
 DB 61 QAAARQKFKSKPRNSDCFOEEDLRQGFQWRKSLPFGAASSTYLNLEKGBGSYATVYVG 120
 QY 121 ISRINGQVALKAVISNNAEEGVPTAIRASLLKGLKHNIVLHDIHTKETLTFVPEY 180
 DB 121 ISRINGQVALKAVISNNAEEGVPTAIRASLLKGLKHNIVLHDIHTKETLTFVPEY 180
 QY 181 MHTDLAQYNSQHPGGLPHNVLRFMFQQLRGLAYIHQVHLRDLKPNQLLSHIGELKL 240
 DB 181 M-----LRFQQLRGLAYIHQVHLRDLKPNQLLSHIGELKL 219
 QY 241 ADFGLARAKSISQYSSSEVTLWYRPPDALLGATYSESLDIWGAGCIFIEMFOQPLF 300
 DB 241 ADFGLARAKSISQYSSSEVTLWYRPPDALLGATYSESLDIWGAGCIFIEMFOQPLF 279
 QY 301 PGVSNILEQLEKIMEVLGVPTEDTWPGVSKLPNYP 336
 DB 301 PGVSNILEQLEKIMEVLGVPTEDTWPGVSKLPNYP 315

RESULT 5
 PFT1_MOUSE STANDARD; PRT; 469 AA.
 ID PFT1_MOUSE
 AC O35495; O35848;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Serine/threonine-protein kinase PFTAIR-1 (EC 2.7.1.37).
 GN Name=Pftk1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, TISSUE
 RP SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RC TISSUE=Testis;
 RX MEDLINE=98208722; PubMed=9547506;
 RA Beszet V., Rhee K., Wolgemuth D.J.;
 RT "The identification and characterization of expression of Pftaire-1, a
 RT novel Cdk family member, suggest its function in the mouse testis and
 RT nervous system.";
 RT Mol. Reprod. Dev. 50:18-29(1998).
 RN [2]
 RP SEQUENCE OF 47-469 FROM N.A., SUBCELLULAR LOCATION, AND DEVELOPMENTAL
 RP STAGE.
 RC STRAIN=NIH Swiss; TISSUE=Brain;
 RX MEDLINE=97345837; PubMed=9202329;
 RA Lazzaro M.A., Albert P.R., Julien J.-P.;
 RT "A novel cdc2-related protein kinase expressed in the nervous
 RT system.";
 RT J. Neurochem. 69:348-364(1997).
 RL [1]
 CC -1- FUNCTION: May play a role in meiosis as well as in neuron
 CC differentiation and/or function.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
 CC -1- TISSUE SPECIFICITY: In the adult, widely expressed at low levels
 CC except in brain, kidney and testis where expression is high. In
 CC the brain, detected in cortex, hippocampus, dentate gyrus, amgdala

```
CC corex, parabuculum and cerebellum. In the embryo, expressed
CC predominantly in the nervous system.
CC -1- DEVELOPMENTAL STAGE: In the testis, expressed at low levels in
CC Sertoli cells of 7-day-old mice, barely detected at day 17, and
CC detected at much higher levels in late pachytene/diplotene
CC spermatocytes in the adult. In the nervous system, expressed at
CC highest levels in the adult.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC CDC2/CDCX subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL: AF033655; AAB87504.1; -.
DR HSSP: P24941; 1G1I.
DR MGD: MGI:894318; Pfck1.
DR InterPro: IPR011009; Kinase_like.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_kin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Nuclear protein; Serine/threonine-protein kinase;
KW Transferrase.
FT DOMAIN 135 419 Protein kinase.
FT NP_BIND 141 149 ATP (By similarity).
FT BINDING 164 164 ATP (By similarity).
FT ACT_SITE 256 256 Proton acceptor (By similarity).
FT CONFLICT 222 223 EQ -> DK (in Ref. 2).
FT CONFLICT 375 375 S -> N (in Ref. 2).
SQ SEQUENCE 469 AA; 53009 MW; 7542DF46D55AC97 CRC64;

Query Match 51.7%; Score 1198.5; DB 1; Length 469;
Best Local Similarity 55.3%; Pred. No. 2e-77;
Matches 246; Conservative 55; Mismatches 103; Indels 41; Gaps 5;

QY 23 GEANSCRSQPET-TEAFKLTDLKEASGSMTSFHRGLQA-----ARA 65
DB 14 GKMKKRLRRLISBSFRIALKEEDTTFDEICVTGSTRNCGTDSVIKHLDTIPEDKKAV 73
QY 66 QRFKS-----KRPFS-NSDCFOBEDLRQGFQNRK-----SLPFGAS 101
DB 74 QRTQSTFDFEKPANQKRVHSENNACINFKSSAKESPKVRRHSSPSSPTSPKFGKD 133
QY 102 SYLNLKLEKSGSAYATYKGISRINQOLVALKVISMNAEKGVPTPTAIRASLKLGLGHANI 161
DB 134 SYEKLEKLEGGSAITYKKGSKNGKLVALKVIRLDEEBCTPTAIRASLKLGLGHANI 193
QY 162 VLLHDIHTKETTLLFVEFYMTDLDLAQYMSQHPGILPHNVRLFMFOLLRLGLAYIHQHYL 221
DB 194 VLLHDIHTKETTLLFVEFYMTDLDLAQYMSQHPGILPHNVRLFMFOLLRLGLAYIHQHYL 253
QY 222 HEDLKQNLILSHLGLKADPGLARAKSIPSTYSSEVVTLMYRPPDALGATVSSSEL 281
DB 254 HEDLKQNLILSHLGLKADPGLARAKSIPSTYSSEVVTLMYRPPDALGATVSSSEL 313
QY 282 DIMGAGCIFIEMFQGPFLPGVSNILEQLEKIWEVGVPTEDTPGVSKLPNNPMPFL 341
DB 314 DMWGVCICIVEMIGVAAFPKMDIDQDLERIFLVLTGTNEDTWPGVSHLPHKPKRFTV 373
QY 342 PTPRSILVWVNRIGRVLEADLASQMLKGFPRDVSQAQALVHYDYSALPSQLYQLPDEE 401
DB 374 YSSKSLRQAMNKLISYVNHAEIDLASKLQSPKRLISAQALSHVEYSDLPRLMELTDMG 433
```

```
QY 402 SLFTVSGVRLKEPKMCDLASYOKH 426
DB 434 SLFTVNVRLQPEAGSKRPAFKNN 458

RESULT 6
ID Q6NVF8 PRELIMINARY; PRT; 423 AA.
AC Q6NVF8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Pfck1 protein.
GN Name=pfck1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lounellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.W., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Jones S.J., Maiz M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.,
RA submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL: BC068134; AAB68134.1; -.
DR InterPro: IPR011199; Cyt_B5_.
DR InterPro: IPR011009; Kinase_like.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_kin_AS.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00191; CYTOCHROME_B5_1; UNKNOWN_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferrase.
SQ SEQUENCE 423 AA; 47684 MW; 5895A066A7C17AAB CRC64;

Query Match 51.4%; Score 1192.5; DB 2; Length 423;
Best Local Similarity 60.2%; Pred. No. 4.7e-77;
Matches 233; Conservative 56; Mismatches 83; Indels 15; Gaps 2;

QY 51 SMTSFHPRGLQAAARQKFSKRPNSDCFOBEDLRQGFQNRK-----SLPFGA 99
```

DB 30 TQSTDPPEKXANQYKRVHSE-----NNACINFKSSAGKESPKVRRHSSPSSPTKPKGX 85
 QY 100 ASSYLNLEKLGEGSAYATYKGISRINGOLVALKVISMAAEGVPTAIRASLKGKXA 159
 DB 86 ADSYKLEKLGEGSAYATYKGISRINGOLVALKVISMAAEGVPTAIRASLKGKXA 145
 QY 160 NIVLHDIHHTKETITLTFPEVYHTDLAQYMSQHPGSLPHNVRLPMFOLLRLAYIHQH 219
 DB 146 NIVLHDIHHTKETITLTFPEVYHTDLAQYMSQHPGSLPHNVRLPMFOLLRLAYIHQH 205
 QY 220 VLRDLKQONLLISHGELKLDLADFGIARAKSIPQSYSEVVTLYWRPPDALLGATEYSS 279
 DB 206 ILHRDLKQONLLISHGELKLDLADFGIARAKSIPQSYSEVVTLYWRPPDALLGATEYSS 265
 QY 280 ELDIWGACCIPIEMFQGPPLFPVGSNILEQLEKIWEVLGVPEDTWPVSKLPNNPEWF 339
 DB 266 CLDMWGVGCIPIEMFQGPPLFPVGSNILEQLEKIWEVLGVPEDTWPVSKLPNNPEWF 325
 QY 340 PLPTPSLHVNNRLGRVPEADLASOMLKGPFRDVSNOELVNDYFSGALPSQYOLPD 399
 DB 326 TVSSKSLROANNKLSYNNHAEADLASOMLKGPFRDVSNOELVNDYFSGALPSQYOLPD 385
 QY 400 ESLEFTVSGVRLKPEKCDLLASQYKQH 426
 DB 386 MSSITFVNVRLQPEAGSMRAFQGN 412

RESULT 7
 AAH68134 PRELIMINARY; PRT; 423 AA.

ID AAH68134 PRELIMINARY; PRT; 423 AA.
 AC AAH68134; PRELIMINARY; PRT; 423 AA.
 DT 14-APR-2004 (TREMBlrel. 27, Created)
 DT 14-APR-2004 (TREMBlrel. 27, Last sequence update)
 DT 14-APR-2004 (TREMBlrel. 27, Last annotation update)
 DE Pfck1 protein.
 GN PFCK1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Boeak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RA Strausberg R.L.,
 RA Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC068134; AAH68134.1; -
 SO SEQUENCE 423 AA; 47668 MW; 5895A06A7C17AAB CRC64;
 Query Match 51.4%; Score 1192.5; DB 2; Length 423;

Best Local Similarity 60.2%; Pred. No. 4,7e-77;
 Matches 233; Conservative 56; Mismatches 83; Indels 15; Gaps 2;
 QY 51 SMTSFPRLGLAARQAKRSKRPNSNDQFOEDIRQGFQWRK-----SLPGA 99
 DB 30 TQSTDPPEKXANQYKRVHSE-----NNACINFKSSAGKESPKVRRHSSPSSPTKPKGX 85
 QY 100 ASSYLNLEKLGEGSAYATYKGISRINGOLVALKVISMAAEGVPTAIRASLKGKXA 159
 DB 86 ADSYKLEKLGEGSAYATYKGISRINGOLVALKVISMAAEGVPTAIRASLKGKXA 145
 QY 160 NIVLHDIHHTKETITLTFPEVYHTDLAQYMSQHPGSLPHNVRLPMFOLLRLAYIHQH 219
 DB 146 NIVLHDIHHTKETITLTFPEVYHTDLAQYMSQHPGSLPHNVRLPMFOLLRLAYIHQH 205
 QY 220 VLRDLKQONLLISHGELKLDLADFGIARAKSIPQSYSEVVTLYWRPPDALLGATEYSS 279
 DB 206 ILHRDLKQONLLISHGELKLDLADFGIARAKSIPQSYSEVVTLYWRPPDALLGATEYSS 265
 QY 280 ELDIWGACCIPIEMFQGPPLFPVGSNILEQLEKIWEVLGVPEDTWPVSKLPNNPEWF 339
 DB 266 CLDMWGVGCIPIEMFQGPPLFPVGSNILEQLEKIWEVLGVPEDTWPVSKLPNNPEWF 325
 QY 340 PLPTPSLHVNNRLGRVPEADLASOMLKGPFRDVSNOELVNDYFSGALPSQYOLPD 399
 DB 326 TVSSKSLROANNKLSYNNHAEADLASOMLKGPFRDVSNOELVNDYFSGALPSQYOLPD 385
 QY 400 ESLEFTVSGVRLKPEKCDLLASQYKQH 426
 DB 386 MSSITFVNVRLQPEAGSMRAFQGN 412

RESULT 8
 ID 06Z037 PRELIMINARY; PRT; 453 AA.
 AC 06Z037;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE MKIA0834 protein (Fragment).
 GN Name=MKIA0834;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryonic tail;
 RX PubMed=14621295;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
 RA Suga Y., Nagase T., Ohara O., Koga H.,
 RA "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT III. the complete nucleotide sequences of 500 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RT DNA Res. 10:167-180(2003).
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AK129226; BAC98036.1; -
 DR InterPro: IPR001199; Cyt. B5.
 DR InterPro: IPR001009; Kinase like.
 DR InterPro: IPR000719; Prot. kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR008271; Ser_thr_pkin_AS.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00069; Kinase; 1.
 DR Prodom: PD000001; Prot. kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00191; CYTOCHROME_B5_1; UNKNOWN_1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KM ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.

```

FT  NON TER      1      1      4DEEB7BAAB23941D CRC64;
SQ  SEQUENCE      453 AA; 51210 MW;
Query Match      51.4%; Score 1192.5; DB 2; Length 453;
Best Local Similarity 60.2%; Pred. No. 5,1e-77;
Matches 233; Conservative 56; Mismatches 83; Indels 15; Gaps 2;

QY  51 SMTSPHRLQARAKQKFKSRPRNSDCFOEEDLRQGFQMK-----SLPFGA 99
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  60 TOSTDPPEKPNQVRRVHSE-----NNACINFKSSAGKESPKVRHSSPSSPTSPKFGK 115
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  100 ASSYLNLEKLGSGSYATYKGISRINGQVALKVISMNAEBGVFPFAIREASLKGELKKA 159
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  116 ADSYEKLEKLGSGSYATYKGISRINGQVALKVISMNAEBGVFPFAIREASLKGELKKA 175
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  160 NIVLHDDIHTKETLTLPVEYVHTDLAQYMSQHPGLPHNVRLPMFOLLRLGLAYIHQH 219
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  176 NIVLHDDIHTKETLTLPVEYVHTDLCOYMDKHPGLHPDNVKLFLFOLLRLGLSYIHQR 235
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  220 VILHRLKPNLLISHLGEIKLADFGIARAKSIPSQTYSSEVYTLWYRPDALLGATEYS 279
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  236 ILHRLKPNLLISDTGELKADFGIARAKSVSHSTYSNEVYTLWYRPDVLGSTEYST 295
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  280 ELDINGAGCIFIEMPOGQPLPGVSNILLEQLEKIMEVLGVPTEDTWPGVSKLPNYNPEWF 339
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  296 CIDMNGVGCIFVEMIQGAFAFPQMDIODQLERITVLGTPEMDTWPGVHSLPHKPERF 355
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  340 PLPTPRSLHVVNRLGRVPEADLASQMLKGFPRDRVSAQOALVHDYFSAIPSOYLQPLD 399
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  356 TVSSSKSLQAVNKNLSYVNHAEIDLASKLQCSPKRRLSAQALSHYFSDLPRLMELTD 415
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  400 EESLFTVSGVRLKPEMCDLLASYOKH 426
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  416 MSSIFTVPVNRLOPEAGSMRAFGKN 442
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
BAC98036      PRELIMINARY; PRT; 453 AA.
AC  BAC98036;
DT  02-MAR-2004 (TREMBLrel. 27, Created)
DT  02-MAR-2004 (TREMBLrel. 27, Last sequence update)
DT  02-MAR-2004 (TREMBLrel. 27, Last annotation update)
DE  MKIAA0034 protein (Fragment).
GN  MKIAA0034.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=embryonic tail;
RA  Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA  Saga Y., Nagase T., Ohara O., Koga H.;
RT  "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
RT  III. The Complete Nucleotide Sequences of 500 Mouse KIAA-homologous
RT  cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT  Randomly Sampled from Size-fractionated Libraries.";
RL  DNA Rep. 10:167-180 (2003).
DR  EMBL; AK129226; BAC98036.1; -.
FT  NON TER      1
SQ  SEQUENCE      453 AA; 51210 MW; 4DEEB7BAAB23941D CRC64;

Query Match      51.4%; Score 1192.5; DB 2; Length 453;
Best Local Similarity 60.2%; Pred. No. 5,1e-77;
Matches 233; Conservative 56; Mismatches 83; Indels 15; Gaps 2;

QY  51 SMTSPHRLQARAKQKFKSRPRNSDCFOEEDLRQGFQMK-----SLPFGA 99
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  60 TOSTDPPEKPNQVRRVHSE-----NNACINFKSSAGKESPKVRHSSPSSPTSPKFGK 115
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  100 ASSYLNLEKLGSGSYATYKGISRINGQVALKVISMNAEBGVFPFAIREASLKGELKKA 159
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

DB  116 ADSYEKLEKLGSGSYATYKGISRINGQVALKVISMNAEBGVFPFAIREASLKGELKKA 175
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  160 NIVLHDDIHTKETLTLPVEYVHTDLAQYMSQHPGLPHNVRLPMFOLLRLGLAYIHQH 219
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  176 NIVLHDDIHTKETLTLPVEYVHTDLCOYMDKHPGLHPDNVKLFLFOLLRLGLSYIHQR 235
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  220 VILHRLKPNLLISHLGEIKLADFGIARAKSIPSQTYSSEVYTLWYRPDALLGATEYS 279
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  236 ILHRLKPNLLISDTGELKADFGIARAKSVSHSTYSNEVYTLWYRPDVLGSTEYST 295
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  280 ELDINGAGCIFIEMPOGQPLPGVSNILLEQLEKIMEVLGVPTEDTWPGVSKLPNYNPEWF 339
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  296 CIDMNGVGCIFVEMIQGAFAFPQMDIODQLERITVLGTPEMDTWPGVHSLPHKPERF 355
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  340 PLPTPRSLHVVNRLGRVPEADLASQMLKGFPRDRVSAQOALVHDYFSAIPSOYLQPLD 399
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  356 TVSSSKSLQAVNKNLSYVNHAEIDLASKLQCSPKRRLSAQALSHYFSDLPRLMELTD 415
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  400 EESLFTVSGVRLKPEMCDLLASYOKH 426
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB  416 MSSIFTVPVNRLOPEAGSMRAFGKN 442
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
Q6DUM7      PRELIMINARY; PRT; 435 AA.
ID  Q6DUM7;
AC  Q6DUM7;
DT  01-OCT-2004 (TREMBLrel. 28, Created)
DT  01-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT  01-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE  Hypothetical protein.
OS  Xenopus laevis (African clawed frog).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC  Xenopodinae; Xenopus.
OX  NCBI_TaxID=8355;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Kidney;
RX  MEDLINE=2238257; PubMed=12477932;
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA  Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA  Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA  Raha S.S., Lochellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA  Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA  Villalon D.K., Wuzny D.M., Sodergren E.U., Lu X., Gibbs R.A.,
RA  Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA  Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA  Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA  Rodriguez A.C., Gimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
RA  Krzywiak M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
RA  Jones S.J., Marra M.A.;
RT  "Generation and initial analysis of more than 15,000 full-length human
RT  and mouse cDNA sequences.";
RL  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Kidney;
RX  MEDLINE=22341132; PubMed=12454917;
RA  Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA  Richardson P.;
RT  "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT  initiative.";
RL  Dev. Dyn. 225:384-391 (2002).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Kidney;
RA  Klein S., Strausberg R.;

```



```
Db 194 VLLHDIHTKETLLVFEVYHTDLCOYMDKHFGJHPDVNKLFLPQLRLGSLYIHQRYIL 253
Qy 222 HNDLKRONLLISHGLKTLADFGILARAKSIPSOYSSSEVVTLMYPPDALLGATEYSSSL 281
Db 254 HNDLKRONLLISDTGSLKTLADFGILARAKSPSHITTSNEVVTLMYPPDVLSTSTYSL 313
Qy 282 DIWGACIETIEMFQGPPLFPGVSNILQLEKIMVELVGPTEDTWPGVSKLPYNPDEWPEPL 341
Db 314 DMWGVCITIVEMIQVAAPFGMKDIDQLEKIFLVLTGTENEDTWPGVSHLPFKPERFTL 373
Qy 342 PTPRSILHVWNRIGRVPEABDLASQMLKPPRDVSAQALVHDYSALPSSQLYQDPDE 401
Db 374 YSSKNI.ROMWNKLSYVNHAEIDLASKLQCSPKNRLSAQALESHEFYSDLPPLMELTWS 433
Qy 402 SLFTVSGVRLKPEMCDLLASYOKGH 426
Db 434 SLFTVENVRLQPEAGSMAFGKN 458
```

RESULT 12

```
08MZ44 PRELIMINARY; PRT; 367 AA.
ID 08MZ44
AC 08MZ44;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE L027880P.
GN Name=Elp63B;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guatin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceolner S.;
RA Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AY113367; AM29372.1; -.
DR HSSP; P24941; 1P2A.
DR FLYBase; FBgn0005640; Elp63B.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Cyt B5.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00191; CYTOCHROME_B5_1; UNKNOWN_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 367 AA; 42309 MW; 05EB51133830B5 CRC64;
```

```
Query Match 46.5%; Score 1078.5; DB 2; Length 367;
Best Local Similarity 60.2%; Pred. No. 5,7e-69;
Matches 213; Conservative 42; Mismatches 84; Indels 15; Gaps 3;
Qy 71 KRPRNSDGF---QEDDLRQGFQWRKSL-----PFGAASSYLNLKLGSGSYATVYKG 120
Db 13 KPRPRKSEVFLNKQETHPR-----RKPRSAFGGDSFPFGQEAAYVVLKPLGSGSYATVYKG 67
```

```
Qy 121 ISRLNGQVLAKVISNNAEGVPTTAIRASLLKGLKHANIVLHDIHTKETLLVFEVY 180
Db 68 FSKLTYYRVALKKEIRLOEEBAGPPTAIRASLLKGLHSHNIVLHDIHTKETLLVFEVY 127
Qy 181 MHTDLAQYMSOHPGGLHPHVRLFLPQLRLGSLYIHHQVHLRDLKRONLLISHGLKTL 240
Db 128 VNTDLSQYMEHHPGGLDHRVRLFLPQLRLGSLYCHRRVVLHROVKRONLLISDGLKTL 187
Qy 241 ADFGIARAKSIPSOYSSSEVVTLMYPPDALLGATEYSSSLDIWGACIETIEMFQGPPLF 300
Db 188 ADFGIARAKSVSHYSHYSEVVTLMYRPDVLGSTEYSTSLDMWGVCITIVEMIQVMTPTF 247
Qy 301 PGVSNILQLEKIMVELVGPTEDTWPGVSKLPYNPDEWPEPLPTPRSLHVWNRIGRVPEA 360
Db 248 PGIHRTYDQDKIRKLTGTPTEDTWPGVTHFPYKPKHKLGYRRKRGHNPRRLYDIIEG 307
Qy 361 EDLASQMLKGPDRDVSAGEALVHDYSALPSSQLYQDPDESLFTVSGVRLKPE 414
Db 308 ETIANGFLQLNPEQRLGADLALQHPYFAQLPKLYELPDFTSIPTVEGVQLYTE 361
```

RESULT 13

```
07KM04 PRELIMINARY; PRT; 501 AA.
ID 07KM04
AC 07KM04;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Serine/threonine protein kinase variant L63B3.
GN Name=L63;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stowers R.S., Garza D., Hogness D.S.;
RA Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AF152403; AAD4514.1; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR InterPro; IPR001199; Cyt B5.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00219; TYKC; 1.
DR PROSITE; PS00191; CYTOCHROME_B5_1; UNKNOWN_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 501 AA; 57741 MW; 0FB0F78DFPFA6DFD CRC64;
```

```
Query Match 46.5%; Score 1078.5; DB 2; Length 501;
Best Local Similarity 60.2%; Pred. No. 8,4e-69;
Matches 213; Conservative 42; Mismatches 84; Indels 15; Gaps 3;
Qy 71 KRPRNSDGF---QEDDLRQGFQWRKSL-----PFGAASSYLNLKLGSGSYATVYKG 120
Db 147 KPRPRKSEVFLNKQETHPR-----RKPRSAFGGDSFPFGQEAAYVVLKPLGSGSYATVYKG 201
Qy 121 ISRLNGQVLAKVISNNAEGVPTTAIRASLLKGLKHANIVLHDIHTKETLLVFEVY 180
Db 202 FSKLTYYRVALKKEIRLOEEBAGPPTAIRASLLKGLHSHNIVLHDIHTKETLLVFEVY 261
```

```

QY 181 MHTDLAQVMSQHPGGLHPHNVRLFMFOLLRGLAYTHHGVHLHPDLKRONLLSHLGLTL 240
DB 262 VNTDLSQYWEKHPGGGLDHRNVRLFLFOLLRGLSLCYKRRVRLHNDVLRPNLLISDCGLTL 321
QY 241 ADFGLARAKSVSHSYSSSEVTLTMYPRLPALGATEYSSELIDWAGCCTFVEMVTCMPTE 300
DB 322 ADFGLARAKSVSHSYSSSEVTLTMYPRLPALGATEYSSELIDWAGCCTFVEMVTCMPTE 381
QY 301 PGVSNILEQLEKIWEVLGVPTEDTWPVGSKLPNYNPEWFPPLPTPSRLHVVNRLGRVPA 360
DB 382 PGIRDTYDQDLKIFLLGLPTEDTWPVGHFPGYKPHKLGFRVPRLLGNHPRLYDIIIG 441

QY 361 EDLASOMLKGPRDRVSAQDALVHDYFSALPSQLYOLPDEBSLFTYSGVRLAKE 414
DB 442 ETIANGFLQNLNPEORLGADDAHQHYPFALPKKLYELPDETSIFVTEGVOLYTE 495

RESULT 14
ID Q7KM08 PRELIMINARY; PRT; 509 AA.
AC Q7KM08;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Serine/threonine protein kinase variant L63A2.
GN Name=L63;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stowers R.S., Garza D., Hogness D.S.;
RA Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL: AF152339; AAD4510.1;
DR GO: GO:0016301; F:kinase activity; IEA.
DR GO: GO:004674; F:protein serine/threonine kinase activity; IEA.
DR InterPro: IPR001189; Cyt B5.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002280; Ser_Thr_kinase.
DR InterPro: IPR008271; Ser_Thr_kin_AS.
DR Pfam: PF00069; Pkinase; 1.
DR Prodom: PD000001; Prot kinase; 1.
DR SMART: SM00220; S_TKc_1.
DR SMART: SM00219; TYKc_1.
DR PROSITE: PS00191; CYTOCHROME B5_1; UNKNOWN_1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS0011; PROTEIN KINASE DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_SF; 1.
KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 509 AA; 58342 MW; F99C2860C8F8FA12A CRC64;

Query Match 46.5%; Score 1078.5; DB 2; Length 509;
Best Local Similarity 60.2%; Pred. No. 8.5e-69;
Matches 213; Conservative 42; Mismatches 84; Indels 15; Gaps 3;

QY 71 KPRPSNSDGF---QEEDLRQGRWRKSL-----PFGAASLYLNEKLGEGSVATVYKG 120
DB 155 KPRPRSEVFLNKQETHPR-----RRRFGAFGDSFPFGQEAIVKLEPREGSVATVYKG 209
QY 121 ISRINQALVALKYISNNAEGVPTAIRASILLKGLKHANIVYLHDIHTKETLTFVFEY 180
DB 210 FSKLTQYRALVKEIRIQEEBGAPFTAIRASLSLKEKHSNIVYLHDIHTRETITVFEY 269
QY 181 MHTDLAQVMSQHPGGLHPHNVRLFMFOLLRGLAYTHHGVHLHPDLKRONLLSHLGLTL 240
DB 270 VNTDLSQYWEKHPGGGLDHRNVRLFLFOLLRGLSLCYKRRVRLHNDVLRPNLLISDCGLTL 329
QY 241 ADFGLARAKSVSHSYSSSEVTLTMYPRLPALGATEYSSELIDWAGCCTFVEMVTCMPTE 300

```

```

DB 330 ADFGLARAKSVSHSYSSSEVTLTMYPRLPALGATEYSSELIDWAGCCTFVEMVTCMPTE 389
QY 301 PGVSNILEQLEKIWEVLGVPTEDTWPVGSKLPNYNPEWFPPLPTPSRLHVVNRLGRVPA 360
DB 390 PGIRDTYDQDLKIFLLGLPTEDTWPVGHFPGYKPHKLGFRVPRLLGNHPRLYDIIIG 449
QY 361 EDLASOMLKGPRDRVSAQDALVHDYFSALPSQLYOLPDEBSLFTYSGVRLAKE 414
DB 450 ETIANGFLQNLNPEORLGADDAHQHYPFALPKKLYELPDETSIFVTEGVOLYTE 503

RESULT 15
ID Q9VZ8 PRELIMINARY; PRT; 509 AA.
AC Q9VZ8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE CG10579-PA (Cg10579-pb) (Lb15250p) (Serine/threonine protein kinase variant L63B2).
GN Name=Elp63E; Synonyms=L63; ORFNames=CG10579;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutcliffe G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazell R.G., Champagne M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabori G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Frankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benson P.V., Betman B.P., Bhargava D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockett P., Broctier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Chertov J.M., Cawley S., Dahlke C., Davoport L.B., Davies P.,
RA de Padlos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garh N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hosten D., Houston K.A., Howland T.J., Wei M.H., Idegawa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Lasbo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matel B., McInosh T.C., McLeod M.P., McPherson D.,
RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshirei A.,
RA Mount S.M., Moy W., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nussekm D.R., Paclet J.M.,
RA Palazozo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodgett J.R., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,
RA Zhang R.A., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RA science 287:2185-2195 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=24246065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frisoe E., Hodgson A.,

```

RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
 RA Paclet J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svitlska R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
 RT "Fishing a whole-genome shotgun: release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminler J.S., Bergman C.M., Krommiller B., Carlson J., Svitlska R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomic perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminler J.S., Milburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Belencourt B.R., Celniker S.E., de Grey A.D., Dysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RG FLYBASE;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RG FLYBASE;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guartin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Munoo J., Paclet J., Paragas V., Park S.,
 RA Patel S., Pounenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RA Stowers R.S., Garza D., Hogness D.S.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AE003478; AAF47781.2; -;
 DR EMBL; BT010064; AAO22533.1; -;
 DR EMBL; AF152402; AAD45513.1; -;
 DR HSSP; P24941; 1P2A.
 DR FLYBASE; FBgn0005640; Eip63E.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR001199; Cyt B5.
 DR InterPro; IPR011009; Kinase_like.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_kin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00191; CYTOCHROME_B5_1; UNKNOWN_1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

SEQ SEQUENCE 509 AA; 58517 MW; 890EFBDC04319DBA CRC64;
 Query Match 46.5%; Score 1078.5; DB 2; Length 509;
 Best Local Similarity 60.2%; Pred. No. 8,5e-69;
 Matches 213; Conservative 42; Mismatches 84; Indels 15; Gaps 3;
 QY 71 KPPRNSDCF--QEDLRQGFQWRKSL-----PFGAASLYNLKLGEGSYATVYKG 120
 DB 155 KPPRPSSEVFLNKQETHR-----RKPSARFGDSPPFGKQEAIVVLEPLGEGSYATVYKG 209
 QY 121 ISRINGQVALLKVISMAEBGVFTAIRBASLLKGLKIANIVLLDIIHTKETLTFVEY 180
 DB 220 PSLTYQVVALKEIRLQEEBGAFTAIRBASLLKGLKKSNIYTLHDIYHTRETLTFVEY 269
 QY 181 MATDIAQYMSQHPGGLHPNNVRLFMFOLLRLGAIYTHQGVLRDLKPNLLISHLGELKL 240
 DB 270 VNTDLSQYWEKIPGGIDHRNVRLFLFOLLRLGISYCHKRRVLRDVKPNLLISDGLKL 329
 QY 241 ADPGLARAKSIPQYSSSEVVTLMYRPPDALLGATEYSSSELDIYGAGCIFEMFGQPLF 300
 DB 330 ADPGLARAKSVPSHTYSHVVTLMYRPPDVLGSTEISTLDMGVGCI FVEMVTGMPTF 389
 QY 301 PGVSNILBQLEKIMVGLVPTEDTPWGVSKLPNYPNPEWFPPTPRSLHVWNNRLGRVPEA 360
 DB 390 PGIIRTYQLDKIFLLGPTPEDTPGVTHFGYKPKHLGFRPRKLGHNFPRLYDIIEG 449
 QY 361 EPLASQMLKGFPRDVSQOELVHDYFSAIPLSQLYQLDDESLFTVSGVRLKPE 414
 DB 450 EYIANGFLQINBEQRIADALQHPYFQLPKLYELPDETSIFVEGVQLYTE 503

Search completed: December 27, 2004, 12:18:33
 Job time : 80 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 26, 2004, 18:49:09 ; Search time 73 Seconds
(without alignment)

2137.633 Million cell updates/sec

Title: US-10-786-065-5

Perfect score: 2320
Sequence: 1 MGDELCACTVDPGCSCHCS.....CDLASYOKGHHRAQPSKCW 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 35872929 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: geneseqp19808:*

2: geneseqp19908:*

3: geneseqp20008:*

4: geneseqp20018:*

5: geneseqp20028:*

6: geneseqp20038:*

7: geneseqp20038:*

8: geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2320	100.0	435	AA018614	Human Ptc
2	2320	100.0	435	ABR57359	Human ser
3	2320	100.0	435	ABR57361	Human ser
4	2320	100.0	435	ADJ96629	Human cyc
5	2088	90.0	405	ADG31726	Human nov
6	2088	87.8	384	ABR57357	Human ser
7	1784	76.9	343	AA018613	Human ptc
8	1659.5	71.5	321	ADG19729	Human kpp
9	1641.5	70.8	345	ADM16427	Human kpp
10	1432.5	61.7	296	ABP96082	Human pro
11	1260.5	54.3	266	AA018617	Human ptc
12	1260.5	54.3	266	ADJ96629	Human cyc
13	1198.5	51.7	469	ABR57358	Human ser
14	1198.5	51.7	469	ADJ76287	Human ser
15	1195	51.5	430	AA018617	Human ptc
16	1194.5	51.5	423	AA018617	Human ptc
17	1194.5	51.5	450	ADN95586	Human bec
18	1194.5	51.5	451	ADN95586	Human bec
19	1194.5	51.5	451	ADJ75554	Human can
20	1194.5	51.5	451	ADJ75554	Human can
21	1191.5	51.4	423	AA018618	Human ptc
22	1191	51.3	330	AA018618	Human ptc
23	1189.5	51.3	330	AA018618	Human ptc
24	1185.5	51.1	469	AA026433	Human hpr
25	1185.5	51.1	469	ADJ69485	Human hea

26	1044.5	45.0	509	4	ABR57849	Abb57849 Drosophila
27	1044.5	45.0	509	4	ABR67228	Abb67228 Drosophila
28	1044.5	45.0	509	4	ABR67227	Abb67227 Drosophila
29	1000	43.1	245	5	AA018616	AA018616 Human ptc
30	993	42.8	240	5	AA018615	AA018615 Human ptc
31	919.5	39.6	497	4	AA039021	AA039021 Human pol
32	916	39.5	496	4	AA078693	AA078693 Human ptc
33	916	39.5	496	6	AA030635	AA030635 Human ptc
34	916	39.5	496	7	ADG62532	ADG62532 Human ptc
35	916	39.5	496	7	ADG62536	ADG62536 Human ptc
36	916	39.5	496	7	ADG62540	ADG62540 Human ptc
37	916	39.5	496	7	ADG38403	ADG38403 Human ptc
38	916	39.5	496	8	AD019444	AD019444 Human ptc
39	910	39.2	527	3	ABR58782	ABR58782 Human ptc
40	909.5	39.2	330	7	ABR57360	ABR57360 Human ptc
41	909.5	39.2	482	8	ADM87193	ADM87193 Human ptc
42	908	39.1	496	4	AA001186	AA001186 Rat gluta
43	907.5	39.1	461	4	AA001187	AA001187 Rat gluta
44	907.5	39.1	462	7	ADG62530	ADG62530 Rat ptc
45	907.5	39.1	462	7	ADG62534	ADG62534 Rat ptc

ALIGNMENTS

RESULT 1
ID AAO18614 standard; protein; 435 AA.
XX
AC AAO18614;
XX
DT 24-OCT-2002 (first entry)
XX
DE Human Ptfaira family kinase splice form 2.
XX
KW Human, Ptfaira family kinase; kinase; enzyme; testis; brain; cytosolic;
KW uterus endometrium adenocarcinoma; lung fibroblast; splice form;
KW kidney renal cell adenocarcinoma; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200261060-A2.
XX
PD 08-AUG-2002.
XX
PF 17-JAN-2002; 2002WO-US001106.
XX
PR 31-JAN-2001; 2001US-0265151P.
PR 09-MAR-2001; 2001US-00801861.
XX
(PEKE) PE CORP NY.
XX
PI Yan C, Ketchum K, Di Francesco V, Beasley EM;
XX
DR WPI; 2002-608515/65.
DR N-PSDB; AAL48890, AAL48891.
XX
PT New human kinase peptide with nucleic acid molecule, useful for treating
PT disorders associated with abnormal expression of kinase protein, e.g.
PT adenocarcinoma of uterus or lung, in drug screening assays and
PT pharmacogenomic analysis.
XX
PS Claim 1; Fig 2; 131pp; English.
XX
CC The present invention provides the protein, cDNA and gene sequences of
CC two splice variants of a human Ptfaira family kinase. The sequences are
CC specifically expressed in the human testis, brain, uterus endometrium
CC adenocarcinomas, lung fibroblasts, kidney renal cell adenocarcinomas, and
CC can be used to treat related diseases. The present sequence is splice
CC variant 2 of the invention
XX
SQ Sequence 435 AA;

```

Query Match      100.0%; Score 2320; DB 5; Length 435;
Best Local Similarity 100.0%; Pred. No. 3.4e-223;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGQELCAKTVQPGSCCYHSGEGEHAHSCRRSQPETTEAAFKLTDLKEASCSMTSPFPRGL 60
DB 1 MGQELCAKTVQPGSCCYHSGEGEHAHSCRRSQPETTEAAFKLTDLKEASCSMTSPFPRGL 60
QY 61 QARAPAKFKSKPRNSDCFOEEDLRQGFQWRKSLPFGAASLYLNLEKLGEGSYATVYKG 120
DB 61 QARAPAKFKSKPRNSDCFOEEDLRQGFQWRKSLPFGAASLYLNLEKLGEGSYATVYKG 120
QY 121 ISRINGQVVALKVISNNAEAGVFTAIRASLLKGLKXANIVLHDIHTKETLTFVFEX 180
DB 121 ISRINGQVVALKVISNNAEAGVFTAIRASLLKGLKXANIVLHDIHTKETLTFVFEX 180
QY 181 MHTDLAQVMSQHPGGHAPHNVRLFMFQLRGALAYIHQVHLHRDLKPOLLLSHLGELKL 240
DB 181 MHTDLAQVMSQHPGGHAPHNVRLFMFQLRGALAYIHQVHLHRDLKPOLLLSHLGELKL 240
QY 241 ADFGLARAKSIPQTSSEVVTLMYRPPDALLGATEYSELDIWAGCIFIEMFGQOPLF 300
DB 241 ADFGLARAKSIPQTSSEVVTLMYRPPDALLGATEYSELDIWAGCIFIEMFGQOPLF 300
QY 301 PGVSNILEQLEKIWEVLGVPTEDTWPGVSKLPNYPNEMFPLPTPRSLHVVMNRLGRVPEA 360
DB 301 PGVSNILEQLEKIWEVLGVPTEDTWPGVSKLPNYPNEMFPLPTPRSLHVVMNRLGRVPEA 360
QY 361 EDLASQMLKGFPRDRVSAQEAALVHDYFSAALPSQLYQLPDEESLFTVSGVRLKPEMCDLLA 420
DB 361 EDLASQMLKGFPRDRVSAQEAALVHDYFSAALPSQLYQLPDEESLFTVSGVRLKPEMCDLLA 420
QY 421 STQKGHHPAQFSKCM 435
DB 421 STQKGHHPAQFSKCM 435

RESULT 2
ABR57359
ID ABR57359 standard; protein; 435 AA.
AC ABR57359;
XX
XX 09-SEP-2003 (first entry)
XX
DE Human serine/threonine protein kinase SEQ ID NO:6.
XX
XX Human; serine/threonine protein kinase; kinase; enzyme; cytosolic;
XX antidiabetic; neuroprotective; nocrotic; antiparkinsonian; analgesic;
XX cerebroprotective; antiinflammatory; gastrointestinal; hepatotropic;
XX gynaecological; uteropathic; dermatological; gene therapy; cancer;
XX diabetes; central nervous system disorder; CNS disorder; liver disease;
XX respiratory disorder; chronic obstructive pulmonary disease; stroke;
XX cardiovascular disorder; dermatological disorder; urological disorder;
XX gastrointestinal disease; haematological disorder; Alzheimer's disease;
XX musculoskeletal disorder; reproductive disorder; Parkinson's disease;
XX neuropathic pain.
XX
OS Homo sapiens.
XX
XX MO2003046167-A1.
XX
XX 05-JUN-2003.
XX
XX 26-NOV-2002; 2002WO-EP013268.
XX
XX 27-NOV-2001; 2001US-033131P.
XX
XX (PARR ) BAYER AG.
XX
XX Koehler RH;
XX
XX WPI, 2003-505196/47.

```

```

DR N-PSDB; ACC79968.
XX
XX New polynucleotide encoding a serine/threonine protein kinase
PT polypeptide, useful for diagnosing, preventing or treating diseases
PT associated with serine/threonine protein kinase dysfunction, e.g. cancer
XX or diabetes.
XX
PS Example 3; Fig 6; 196pp; English.
XX
XX The present sequence represents a human serine/threonine protein kinase
CC (1). (1) has cytosolic, antidiabetic, neuroprotective, antiparkinsonian,
CC nocrotic, cerebroprotective, analgesic, antiinflammatory, hepatotropic,
CC gastrointestinal, gynaecological, uteropathic and dermatological
CC activities, and can be used in gene therapy. Serine/threonine protein
CC kinase polynucleotide and polypeptide sequences can be used in
CC diagnosing, preventing, ameliorating or treating diseases associated with
CC serine/threonine protein kinase dysfunction. They may also be used to
CC identify test compounds that may act, for example, as activators or
CC inhibitors at the enzyme's active site. The human serine/threonine
CC protein kinase and its fragments are also useful in raising specific
CC antibodies that can block the enzyme and effectively reduce its activity.
CC Human serine/threonine protein kinase sequences can be used in the
CC preparation of a medicament for modulating the activity of a serine/
CC threonine protein kinase in a disease, e.g. cancer, diabetes, a central
CC nervous system (CNS) disorder, a respiratory disorder (including chronic
CC obstructive pulmonary disease), a cardiovascular disorder, a
CC haematological disorder, a gastrointestinal or liver disease, a
CC neurological disorder, a musculoskeletal disorder, a reproductive
CC disorder, or a urological disorder. CNS disorders may include Alzheimer's
CC disease, Parkinson's disease, stroke or neuropathic pain
XX
SQ Sequence 435 AA;

```

```

Query Match      100.0%; Score 2320; DB 7; Length 435;
Best Local Similarity 100.0%; Pred. No. 3.4e-223;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGQELCAKTVQPGSCCYHSGEGEHAHSCRRSQPETTEAAFKLTDLKEASCSMTSPFPRGL 60
DB 1 MGQELCAKTVQPGSCCYHSGEGEHAHSCRRSQPETTEAAFKLTDLKEASCSMTSPFPRGL 60
QY 61 QARAPAKFKSKPRNSDCFOEEDLRQGFQWRKSLPFGAASLYLNLEKLGEGSYATVYKG 120
DB 61 QARAPAKFKSKPRNSDCFOEEDLRQGFQWRKSLPFGAASLYLNLEKLGEGSYATVYKG 120
QY 121 ISRINGQVVALKVISNNAEAGVFTAIRASLLKGLKXANIVLHDIHTKETLTFVFEX 180
DB 121 ISRINGQVVALKVISNNAEAGVFTAIRASLLKGLKXANIVLHDIHTKETLTFVFEX 180
QY 181 MHTDLAQVMSQHPGGHAPHNVRLFMFQLRGALAYIHQVHLHRDLKPOLLLSHLGELKL 240
DB 181 MHTDLAQVMSQHPGGHAPHNVRLFMFQLRGALAYIHQVHLHRDLKPOLLLSHLGELKL 240
QY 241 ADFGLARAKSIPQTSSEVVTLMYRPPDALLGATEYSELDIWAGCIFIEMFGQOPLF 300
DB 241 ADFGLARAKSIPQTSSEVVTLMYRPPDALLGATEYSELDIWAGCIFIEMFGQOPLF 300
QY 301 PGVSNILEQLEKIWEVLGVPTEDTWPGVSKLPNYPNEMFPLPTPRSLHVVMNRLGRVPEA 360
DB 301 PGVSNILEQLEKIWEVLGVPTEDTWPGVSKLPNYPNEMFPLPTPRSLHVVMNRLGRVPEA 360
QY 361 EDLASQMLKGFPRDRVSAQEAALVHDYFSAALPSQLYQLPDEESLFTVSGVRLKPEMCDLLA 420
DB 361 EDLASQMLKGFPRDRVSAQEAALVHDYFSAALPSQLYQLPDEESLFTVSGVRLKPEMCDLLA 420
QY 421 STQKGHHPAQFSKCM 435
DB 421 STQKGHHPAQFSKCM 435

RESULT 3
ABR57361
ID ABR57361 standard; protein; 435 AA.

```

XX ABR57361;
AC
XX
DT 09-SEP-2003 (first entry)
XX
XX Human serine/threonine protein kinase SEQ ID NO:10.
DE
XX
XX Human; serine/threonine protein kinase; kinase; enzyme; cytosolic;
KW antidiabetic; neuroprotective; nootropic; antiparkinsonian; analgesic;
KW cerebroprotective; antiinflammatory; gastrointestinal; hepatocytic;
KW gynaecological; uropathic; dermatological; gene therapy; cancer;
KW diabetes; central nervous system disorder; CNS disorder; liver disease;
KW respiratory disorder; chronic obstructive pulmonary disease; stroke;
KW cardiovascular disorder; dermatological disorder; urological disorder;
KW gastrointestinal disease; haematological disorder; Alzheimer's disease;
KW musculoskeletal disorder; reproductive disorder; Parkinson's disease;
KW neuropathic pain.
XX
XX Homo sapiens.
OS
XX WO2003046167-A1.
PN
XX
XX 05-JUN-2003.
PD
XX
XX 26-NOV-2002; 2002WO-EP013268.
PP
XX
XX 27-NOV-2001; 2001US-033131P.
PR
XX
XX (FARB) BAYER AG.
PA
XX
XX Koehler RH;
PI
XX
XX WPI; 2003-505196/47.
DR
XX N-PSDB; ACC79970, ACC79971.
XX
XX New polynucleotide encoding a serine/threonine protein kinase
PT polypeptide, useful for diagnosing, preventing or treating diseases
PT associated with serine/threonine protein kinase dysfunction, e.g. cancer
PT or diabetes.
XX
XX
XX Disclosure; Page 188-190; 196pp; English.
XX
XX The present sequence represents a human serine/threonine protein kinase
CC (1). (1) has cytosolic, antidiabetic, neuroprotective, antiparkinsonian,
CC nootropic, cerebroprotective, analgesic, antiinflammatory, hepatocytic,
CC gastrointestinal, gynaecological, uropathic and dermatological
CC activities, and can be used in gene therapy. Serine/threonine protein
CC kinase polynucleotide and polypeptide sequences can be used in
CC diagnosing, preventing, ameliorating or treating diseases associated with
CC serine/threonine protein kinase dysfunction. They may also be used to
CC identify test compounds that may act, for example, as activators or
CC inhibitors at the enzyme's active site. The human serine/threonine
CC protein kinase and its fragments are also useful in raising specific
CC antibodies that can block the enzyme and effectively reduce its activity.
CC Human serine/threonine protein kinase sequences can be used in the
CC preparation of a medicament for modulating the activity of a serine/
CC threonine protein kinase in a disease, e.g. cancer, diabetes, a central
CC nervous system (CNS) disorder, a respiratory disorder (including chronic
CC obstructive pulmonary disease), a cardiovascular disorder, a
CC dermatological disorder, a gastrointestinal or liver disease, a
CC haematological disorder, a musculoskeletal disorder, a reproductive
CC disorder, or a urological disorder. CNS disorders may include Alzheimer's
CC disease, Parkinson's disease, stroke or neuropathic pain
XX
XX
XX Sequence 435 AA;
SQ

Query Match 100.0%; Score 2320; DB 7; Length 435;
Best Local Similarity 100.0%; Pred. No. 3.4e-223;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MGBELCAKTVVOPGSCYHSGEGEASHCRSOPETTEAFAFKLTDLEASGSMSPFPRGL 60
1 MGBELCAKTVVOPGSCYHSGEGEASHCRSOPETTEAFAFKLTDLEASGSMSPFPRGL 60

QY 61 QAAAPQKFKSKRPKRSNSDGFQEDLRQGFQMRKSLPFGAASSYLNIETKIGEGSYATVYKG 120
DB 61 QAAAPQKFKSKRPKRSNSDGFQEDLRQGFQMRKSLPFGAASSYLNIETKIGEGSYATVYKG 120
QY 121 ISRINGQVALKVIYSNNAEEGVPTAIRASLLKGLKHNIVLHDIHTKELTFVFEY 180
DB 121 ISRINGQVALKVIYSNNAEEGVPTAIRASLLKGLKHNIVLHDIHTKELTFVFEY 180
QY 181 MHTDLAQVNSQHPGGLHPNNVLFMFQILRGALVIHQVHLRDLKPQNLISHLGEKL 240
DB 181 MHTDLAQVNSQHPGGLHPNNVLFMFQILRGALVIHQVHLRDLKPQNLISHLGEKL 240
QY 241 ADPGLARAKSISQIYSSSVTVLWTRPPDALLGATRYSESLDWAGACFIEMFGQPLF 300
DB 241 ADPGLARAKSISQIYSSSVTVLWTRPPDALLGATRYSESLDWAGACFIEMFGQPLF 300
QY 301 PGVSNLEQLEKIMETLVGPTEDTWPGVSKLPNNYPENPPLTPPSLHVNNRLGRVPEA 360
DB 301 PGVSNLEQLEKIMETLVGPTEDTWPGVSKLPNNYPENPPLTPPSLHVNNRLGRVPEA 360
QY 361 EDLASQMLKGFPRDRVSAQEALVHDYFSALPQOLYQLPDEESLFTVSGVRLKPEMCDLLA 420
DB 361 EDLASQMLKGFPRDRVSAQEALVHDYFSALPQOLYQLPDEESLFTVSGVRLKPEMCDLLA 420
QY 421 SYOKGHHHPAQFSKCW 435
DB 421 SYOKGHHHPAQFSKCW 435

RESULT 4
ADJ96629
ID ADJ96629 standard; protein; 435 AA.
XX
XX
AC ADJ96629;
XX
DT 06-MAY-2004 (first entry)
XX
XX Human cyclin dependent kinase PFTAIR2 protein SeqID 86.
DE
XX
XX kinase; human; tyrosine protein kinase; serine/threonine protein kinase;
KW kinase; human; tyrosine protein kinase; serine/threonine protein kinase;
KW PTK; gene therapy; cancer; immune-related disease; metabolic;
KW cardiovascular disease; brain; neuronal associated disease; metabolic;
KW inflammatory disorder; cytosolic; neuroprotective; immunomodulator;
XX antiinflammatory; enzyme; cyclin dependent kinase; PFTAIR2.
XX
OS Homo sapiens.
OS
XX
XX WO2004006838-A2.
PN
XX
XX 22-JAN-2004.
PD
XX
XX 15-JUL-2003; 2003WO-US021730.
PP
XX
XX 15-JUL-2002; 2002US-0395632P.
PR
XX
XX (SUGB-) SUGEN INC.
PA
XX
XX whyte D, Manning G, Caenepeel S;
PI
XX
XX WPI; 2004-122753/12.
DR
XX N-PSDB; ADJ96563.
DR
XX
XX New nucleic acid molecule encoding a kinase polypeptide, useful for
PT preparing a composition for treating diseases or disorders, e.g., cancer,
PT or neurological, immunological or inflammatory disorders.
PT
PS Claim 1; SEQ ID NO 86; 366pp; English.
PS
XX
XX This invention relates to a novel isolated, enriched or purified nucleic
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC to human tyrosine and serine/threonine protein kinases (PTK's and STK's),

as well as protein kinase-like enzymes. The present invention describes screening methods to identify agonists, antagonists and antibodies that can be used to modulate the activity or function of the mammalian kinase enzymes. As such, these compositions can be used for gene therapy purposes to treat diseases or disorders including cancer, immune-related diseases, cardiovascular disease, brain or neuronal associated disease, metabolic and inflammatory disorders. Accordingly, they exhibit cytostatic, neuroprotective, immunomodulator and antiinflammatory activities. This polypeptide sequence is a human kinase protein sequence of the invention.

Sequence 435 AA:

Query Match 100.0%; Score 2320; DB 8; Length 435;
Best Local Similarity 100.0%; Pred. No. 3.4e-223;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDELCAKTYOPCSCYHSGEGEASHCRSOPETTEAATKLDLKEASCSMTSFPRGL 60
DB 1 MGDELCAKTYOPCSCYHSGEGEASHCRSOPETTEAATKLDLKEASCSMTSFPRGL 60
QY 61 QARAAQKFKSKRRNSDCFOEBDLKQGFQWRKSLPFGAASLYINLEKLGEGSYATVYKG 120
DB 61 QARAAQKFKSKRRNSDCFOEBDLKQGFQWRKSLPFGAASLYINLEKLGEGSYATVYKG 120
QY 121 ISRINGQVALAKTISNNAEGVPTAIRASLILKGLKHANIVLHDIHTKETLTVEFY 180
DB 121 ISRINGQVALAKTISNNAEGVPTAIRASLILKGLKHANIVLHDIHTKETLTVEFY 180
QY 181 MHTDLAQYMSQHGGLAPHNVRLFMFOLLRLGLAYIHHQVHLHDLKPNMLLSHLELKL 240
DB 181 MHTDLAQYMSQHGGLAPHNVRLFMFOLLRLGLAYIHHQVHLHDLKPNMLLSHLELKL 240
QY 241 ADPGLARAKSIPQYTSSEVVTLMYRPPDALGATEYSSELIDWAGCIFIEMFGQPLF 300
DB 241 ADPGLARAKSIPQYTSSEVVTLMYRPPDALGATEYSSELIDWAGCIFIEMFGQPLF 300
QY 301 PGVSNILEQLEKIEWELGVPTEDTWGVSKLPTNPNBWPFLPTPRSLHVTYNNRLGRVPER 360
DB 301 PGVSNILEQLEKIEWELGVPTEDTWGVSKLPTNPNBWPFLPTPRSLHVTYNNRLGRVPER 360
QY 361 EDLASQMLKGFPRDRVSAQBALVHDYFSALPSQLYQLPDESLFTVSGVRLKPEMCDLLA 420
DB 361 EDLASQMLKGFPRDRVSAQBALVHDYFSALPSQLYQLPDESLFTVSGVRLKPEMCDLLA 420
QY 421 SYQKGHHPAQFSKCM 435
DB 421 SYQKGHHPAQFSKCM 435

RESULT 5
ADCC1726
ID ADCC1726 standard; protein; 405 AA.

ADCC1726;

18-DEC-2003 (first entry)

Human novel polypeptide sequence, SEQ ID NO:1808.

Human; diagnostic; drug screening; forensics; gene mapping;
biodiversity assessment; Parkinson's disease; Alzheimer's disease;
neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
ulcers; osteoporosis; autoimmune disease; cancer;
molecular weight marker; food supplement; antiparkinsonian; nootropic;
neuroprotective; antihaemic; anticoagulant; thrombolytic; vulnerary;
antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
gene therapy.

Homo sapiens.
MO2003029271-AA.

10-APR-2003.

24-SEP-2002; 2002WO-US030474.

24-SEP-2001; 2001US-0324631P.

(HYSE-) HYSEQ INC.

Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

Zhou P, Ghosh M, Wang D, Ma Y, Aundi V, Wang Z, Wang G;

Haley-Vicente D, Drmanac RT;

WPI: 2003-371981/35.

N-PSDB; ADCC30755.

New polynucleotide and polypeptide useful for diagnosing, preventing or treating conditions such as neurodegenerative diseases, anemias, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer.

Claim 20; SEQ ID NO 1808; 1185bp; English.

The invention relates to 971 novel human cDNA sequences (ADCC29919-ADCC30889) and the polypeptides they encode (ADCC30890-ADCC31860). The invention also relates to nucleic acid sequences over 99% identical with the novel human cDNAs. The invention additionally encompasses expression vectors and host cells comprising a nucleic acid of the invention; the recombinant production of a polypeptide of the invention; an antibody against a polypeptide of the invention; a method of detecting polynucleotides or polypeptides of the invention; and methods of identifying a compound which binds to a polypeptide of the invention. The invention further discloses methods of preventing, treating or ameliorating a medical condition; kits comprising polynucleotide probes and/or monoclonal antibodies for carrying out the methods of the invention; method for the identification of compounds that modulate the expression or activity of the polynucleotide and/or polypeptide; and 767 config sequences corresponding to the cDNA sequences of the invention (ADCC1861-ADCC32627) and the polypeptides encoded by the configs (ADCC32628-ADCC33394). The nucleic acids and polypeptides of the invention are useful in diagnostics, drug screening, forensics, gene mapping, in the identification of mutations responsible for genetic disorders or other traits, for assessing biodiversity, and in producing many other types of data and products dependent on DNA and amino acid sequences. They are also used for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, anaemia, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer. The nucleic acids may also be used as hybridisation probes or primers, and in the recombinant production of a protein. The polypeptides are also useful in generating antibodies, as molecular weight markers, and as food supplements. The present sequence represents a specifically claimed human polypeptide sequence of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 405 AA:

Query Match 90.0%; Score 2088; DB 7; Length 405;
Best Local Similarity 100.0%; Pred. No. 5.7e-200;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 LNDLKEASCSMTSFPRGLQARAAQKFKSKRRNSDCFOEBDLKQGFQWRKSLPFGAAS 101
DB 12 LNDLKEASCSMTSFPRGLQARAAQKFKSKRRNSDCFOEBDLKQGFQWRKSLPFGAAS 71
QY 102 STYINLEKLGEGSYATVYKGISRINGQVALAKTISNNAEGVPTAIRASLILKGLKHANI 161
DB 72 STYINLEKLGEGSYATVYKGISRINGQVALAKTISNNAEGVPTAIRASLILKGLKHANI 131
QY 162 VLLHDIHTKETLTVEFYMTDLAQYMSQHGGLAPHNVRLFMFOLLRLGLAYIHHQVHL 221
DB 132 VLLHDIHTKETLTVEFYMTDLAQYMSQHGGLAPHNVRLFMFOLLRLGLAYIHHQVHL 191

QY 222 HRDLKPNLLISHLGLKADFGRLARAKSPSQTSSEVVTLMYRPPDALLGATEYSSL 281
 DB 192 HRDLKPNLLISHLGLKADFGRLARAKSPSQTSSEVVTLMYRPPDALLGATEYSSL 251
 QY 282 DIMGAGCIFIEMFQOQPLPPGVSNILEQLEKIWEVLGVPTEDTWPQVSKLPYNNBEMFPL 341
 DB 252 DIMGAGCIFIEMFQOQPLPPGVSNILEQLEKIWEVLGVPTEDTWPQVSKLPYNNBEMFPL 311
 QY 342 PPRRLHYVVMNRGLGRVPEAEDLASQMLKGFPPDRVSAQALVHDFSAALPSQLYQLPDEE 401
 DB 312 PPRRLHYVVMNRGLGRVPEAEDLASQMLKGFPPDRVSAQALVHDFSAALPSQLYQLPDEE 371
 QY 402 SLFTVSGVRLKPEMCDLLASYOKGHHPPAOFKSCW 435
 DB 372 SLFTVSGVRLKPEMCDLLASYOKGHHPPAOFKSCW 405

RESULT 6
 ABR57357
 ID ABR57357 standard, protein, 384 AA.
 XX ABR57357;
 AC
 XX
 DT 09-SEP-2003 (first entry)
 XX
 DE Human serine/threonine protein kinase SEQ ID NO.2.
 XX
 XX Human; serine/threonine protein kinase; kinase; enzyme; cytosolic;
 KW antidiabetic; neuroprotective; nootropic; antiparkinsonian; analgesic;
 KW cerebroprotective; antiinflammatory; gastrointestinal; hepatotropic;
 KW gynaecological; uropathic; dermatological; gene therapy; cancer;
 KW diabetes; central nervous system disorder; CNS disorder; liver disease;
 KW respiratory disorder; chronic obstructive pulmonary disease; stroke;
 KW cardiovascular disorder; dermatological disorder; urological disorder;
 KW gastrointestinal disease; haematological disorder; Alzheimer's disease;
 KW musculoskeletal disorder; reproductive disorder; Parkinson's disease;
 KW neuropathic pain.
 KW
 OS Homo sapiens.
 XX
 PN WO2003046167-A1.
 XX
 PD 05-JUN-2003.
 XX
 PF 26-NOV-2002; 2002WO-EP013268.
 XX
 PR 27-NOV-2001; 2001US-033131P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Koehler RH;
 DR N-PSDB; ACC79967.
 XX
 PT WPI: 2003-505196/47.
 PT New polynucleotide encoding a serine/threonine protein kinase
 PT polypeptide, useful for diagnosing, preventing or treating diseases
 PT associated with serine/threonine protein kinase dysfunction, e.g. cancer
 PT or diabetes.
 XX
 PS Claim 1; Fig 2; 196bp; English.
 XX
 CC The present sequence represents a human serine/threonine protein kinase
 CC (I). (I) has cytosolic, antidiabetic, neuroprotective, antiparkinsonian,
 CC nootropic, cerebroprotective, analgesic, antiinflammatory, hepatotropic,
 CC gastrointestinal, gynaecological, uropathic and dermatological
 CC activities, and can be used in gene therapy. Serine/threonine protein
 CC kinase polynucleotide and polypeptide sequences can be used in
 CC diagnosing, preventing, ameliorating or treating diseases associated with
 CC serine/threonine protein kinase dysfunction. They may also be used to
 CC identify test compounds that may act, for example, as activators or
 CC inhibitors at the enzyme's active site. The human serine/threonine
 CC protein kinase and its fragments are also useful in raising specific

CC antibodies that can block the enzyme and effectively reduce its activity.
 CC Human serine/threonine protein kinase sequences can be used in the
 CC preparation of a medicament for modulating the activity of a serine/
 CC threonine protein kinase in a disease, e.g. cancer, diabetes, a central
 CC nervous system (CNS) disorder, a respiratory disorder (including chronic
 CC obstructive pulmonary disease), a cardiovascular disorder, a
 CC haematological disorder, a gastrointestinal or liver disease, a
 CC neurological disorder, a musculoskeletal disorder, a reproductive
 CC disorder, or a urological disorder. CNS disorders may include Alzheimer's
 CC disease, Parkinson's disease, stroke or neuropathic pain
 CC XX

Sequence 384 AA;
 SQ

Query Match 87.8%; Score 2038; DB 7; Length 384;
 Best Local Similarity 100.0%; Pred. No. 5.4e-195;
 Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 MTSFPRGLQARARAKSPKSPRSNSDCFOEEDLRQGFQWRKSLPFGAASSYLNLEKGE 111
 DB 1 MTSFPRGLQARARAKSPKSPRSNSDCFOEEDLRQGFQWRKSLPFGAASSYLNLEKGE 60
 QY 112 GSYATVYKGISIRINGQLVALKYISMAEBGVFTAIRASLKLKGNANTVLLHDIHTK 171
 DB 61 GSYATVYKGISIRINGQLVALKYISMAEBGVFTAIRASLKLKGNANTVLLHDIHTK 120
 QY 172 ETLTFVFETMTDLAQVMSQHRCGLHPHNVRLEFMQLRGALYIHQVHLHRDLKPNLL 231
 DB 121 ETLTFVFETMTDLAQVMSQHRCGLHPHNVRLEFMQLRGALYIHQVHLHRDLKPNLL 180
 QY 232 ISHLGELKLDLGLARAKSPQTSSEVVTLMYRPPDALLGATEYSSSLDIMGAGCIFI 291
 DB 181 ISHLGELKLDLGLARAKSPQTSSEVVTLMYRPPDALLGATEYSSSLDIMGAGCIFI 240
 QY 292 EMFQOQPLPPGVSNILEQLEKIWEVLGVPTEDTWPQVSKLPYNNBEMFPLPPRSLHYVM 351
 DB 241 EMFQOQPLPPGVSNILEQLEKIWEVLGVPTEDTWPQVSKLPYNNBEMFPLPPRSLHYVM 300
 QY 352 NRLGRVPEAEDLASQMLKGFPPDRVSAQALVHDFSAALPSQLYQLPDEESSLFTVSGVRL 411
 DB 301 NRLGRVPEAEDLASQMLKGFPPDRVSAQALVHDFSAALPSQLYQLPDEESSLFTVSGVRL 360
 QY 412 KPEMCDLLASYOKGHHPPAOFKSCW 435
 DB 361 KPEMCDLLASYOKGHHPPAOFKSCW 384

RESULT 7
 AAO18613
 ID AAO18613 standard, protein, 343 AA.
 XX
 AC AAO18613;
 XX
 DT 24-OCT-2002 (first entry)
 XX
 DE Human peltaire family kinase splice form 1.
 XX
 XX Human; peltaire family kinase; kinase; enzyme; testis; brain; cytosolic;
 KW uterus endometrium adenocarcinoma; lung fibroblast; splice form;
 KW kidney renal cell adenocarcinoma; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200261060-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 17-JAN-2002; 2002WO-US001106.
 XX
 PR 31-JAN-2001; 2001US-0265151P.
 PR 09-MAR-2001; 2001US-00801861.
 XX
 PA (PEKE) PE CORP NY.

PI Yan C, Ketchum X, Di Francesco V, Beasley EM;
 XX
 DR WPI; 2002-608515/65.
 XX N-PSDB; AAL48889, AAL48890.
 PT New human kinase peptide and nucleic acid molecule, useful for treating
 PT disorders associated with abnormal expression of kinase protein, e.g.
 PT adenocarcinoma of uterus or lung, in drug screening assays and
 PT pharmacogenomic analysis.
 XX
 PS Claim 1; Fig 2; 131pp; English.
 XX
 CC The present invention provides the protein, cDNA and gene sequences of
 CC two splice variants of a human pftaire family kinase. The sequences are
 CC specifically expressed in the human testis, brain, uterus endometrium
 CC adenocarcinoma, lung fibroblasts, kidney renal cell adenocarcinoma, and
 CC can be used to treat related diseases. The present sequence is splice
 CC variant 1 of the invention
 XX
 SQ Sequence 343 AA;
 Query Match 76.3%; Score 1784; DB 5; Length 343;
 Best Local Similarity 100.0%; Pred. No. 1.3e-169;
 Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGELCAKTVQPCSCYHSGEGEHAHSCRSQPEETBAAPKLTDLKASGSMTHPRGL 60
 Db 1 MGELCAKTVQPCSCYHSGEGEHAHSCRSQPEETBAAPKLTDLKASGSMTHPRGL 60
 QY 61 QAAARAQKFKSKPRNSDCQFEEEDLRQGFQWRKSLPFGAASLYLTKLGEGSYATVYKG 120
 Db 61 QAAARAQKFKSKPRNSDCQFEEEDLRQGFQWRKSLPFGAASLYLTKLGEGSYATVYKG 120
 QY 121 ISRLNGQLVALKVIYSMNAEEGVPTAIRESALLKGLKHNIVLLHDIHTKETLTFFVPEX 180
 Db 121 ISRLNGQLVALKVIYSMNAEEGVPTAIRESALLKGLKHNIVLLHDIHTKETLTFFVPEX 180
 QY 181 MHTDLAOWMSQHGGHHPHNVRLFMFQLRGLAYIHQVYLDLKPQULLISHGLK 240
 Db 181 MHTDLAOWMSQHGGHHPHNVRLFMFQLRGLAYIHQVYLDLKPQULLISHGLK 240
 QY 241 ADPGLARAKSIPSQYTSSEVVTLMYRPPDALLGATEYSSSELDIWGAGCTFIEMFGQPLF 300
 Db 241 ADPGLARAKSIPSQYTSSEVVTLMYRPPDALLGATEYSSSELDIWGAGCTFIEMFGQPLF 300
 QY 301 PGVSNILBQLEKIMWEVLGVPTEDTPWGVSKLPYNNP 336
 Db 301 PGVSNILBQLEKIMWEVLGVPTEDTPWGVSKLPYNNP 336
 RESULT 8
 AAG79729
 ID AAG79729 standard; protein; 321 AA.
 AC AAG79729;
 XX
 DT 04-MAR-2003 (first entry)
 XX
 XX Human KPP-7, Incyte ID No. 7494145D01.
 KW Kinase; phosphatase; KPP; cell proliferation; arteriosclerosis;
 KW atherosclerosis; cirrhosis; hepatitis; reproduction; infertility;
 KW paroxysmal nocturnal haemoglobinuria; polycythaemia vera; psoriasis;
 KW primary chromocytopenia; cancer; development; renal tubular acidosis;
 KW anaemia; mental retardation; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; epilepsy; Sjogren's syndrome; uveitis; asthma;
 KW menstrual cycle; autoimmune; inflammation; rheumatoid arthritis; AIDS;
 KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; allergy;
 KW diabetes mellitus; glomerulonephritis; Goodpasture's syndrome; gout;
 KW Graves' disease; Hashimoto's thyroiditis; irritable bowel syndrome;
 KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
 KW Reiter's syndrome; mouse; PFTAIR kinase; enzyme.
 XX

OS Homo sapiens.
 XX
 FH Key
 FT Modified-site
 FT Modified-site
 FT Modified-site
 FT Domain
 FT Domain
 FT Binding-site
 FT Modified-site
 FT Peptide
 FT Active-site
 FT Protein
 FT Modified-site
 FT Domain
 FT Domain
 FT WO200290530-A2.
 PD 14-NOV-2002.
 XX
 PF 16-JAN-2002; 2002MO-US001369.
 XX
 PR 18-JAN-2001; 2001US-0263083P.
 PR 23-FEB-2001; 2001US-0271117P.
 PR 23-FEB-2001; 2001US-0271205P.
 PR 16-MAR-2001; 2001US-0276859P.
 PR 23-MAR-2001; 2001US-0278504P.
 PR 23-MAR-2001; 2001US-0278522P.
 PR 29-MAR-2001; 2001US-0280266P.
 PR 29-MAR-2001; 2001US-0280510P.
 XX
 PA (INCYTE GENOMICS INC.
 XX
 XX Lee EA, Walla NK, Baughn MR, Ison CH, Gururajan R, Arvizu C;
 PI Yao MG, Jackson JL, Tang TY, Yue H, Tran B, Ding L, Lu DM;
 PI Lal PG, Warren BA;
 XX
 DR WPI; 2003-111972/10.
 DR N-PSDB; ABA00722.
 XX
 XX New human kinases and phosphatases and polynucleotides, useful for
 PT diagnosing, treating or preventing autoimmune or inflammatory disorders
 PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
 PT cancer or hepatitis.
 XX
 PS Claim 1; Page 135-36; 144pp; English.
 XX
 CC The sequences given in AAG79723-30 show human kinases and phosphatases
 CC (KPP). The KPP polypeptides, polynucleotides, and agonists and
 CC antagonists to them, are useful for diagnosing, treating or preventing
 CC disorders associated with aberrant expression of KPP, particularly cell
 CC proliferative disorders (e.g. arteriosclerosis, atherosclerosis,
 CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
 CC vera, psoriasis, primary thrombocytopenia or cancer), developmental
 CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
 CC neurological disorders (e.g. Alzheimer's disease, Parkinson's disease or

CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
 CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
 CC allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's
 CC disease, diabetes mellitus, glomerulonephritis, Goodpasture's syndrome,
 CC gout, Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome,
 CC multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's
 CC syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,
 CC bacterial, fungal, parasitic, protozoan or helminthic infections. They
 CC are also useful in the assessment of the effects of exogenous compounds
 CC on the expression of nucleic acid and amino acid sequences of proteins
 CC associated with KRP. This protein is homologous to mouse PFTAIRK kinase
 XX
 XX Sequence 321 AA;

Query Match 71.5%; Score 1659.5; DB 6; Length 321;
 Best Local Similarity 83.6%; Pred. No. 3.7e-157;
 Matches 321; Conservative 0; Mismatches 0; Indels 63; Gaps 1;

QY 52 MTSFPRGQARARQKFKSKPRNSDCFOEDLRGCFQWRKSLPFGAASLYLNLEKLG 111
 DB 1 MTSFPRGQARARQKFKSKPRNSDCFOEDLRGCFQWRKSLPFGAASLYLNLEKLG 60
 QY 112 GSATATYKGISIRINGQLVALKVIISMAEGVPPTAIRASLLKGLKANIYVLLHDIHTK 171
 DB 61 GSATATYKGISIRINGQLVALKVIISMAEGVPPTAIRASLLKGLKANIYVLLHDIHTK 120
 QY 172 ETLTFVFEYMTDLAQYMSQHFGGLHPHNVRLFMQLRLGLAYIHQHVLRDLKPNLL 231
 DB 121 ETLTFVFEYMTDLAQYMSQHFGGLHPHNVRLFMQLRLGLAYIHQHVLRDLKPNLL 180
 QY 232 ISHLGELKADGRLARAKISPSQTSSEVYTLWYRPPDALLGATEYSELDIWGACIT 291
 DB 181 ISHLGELKADGRLARAKISPSQTSSEVYTLWYRPPDALLGATEYSELDIWGACIT 240
 QY 292 EHFQGPPLPGVSNILEQLEKIMEVLTGPTEDTWPVSKLPNNPWFPLPRSLHYVM 351
 DB 241 EHFQGPPLPGVSNILEQLEKIMEVLTGPTEDTWPVSKLPNNPWFPLPRSLHYVM 285
 QY 352 NRLGVRPEADLASQMLKGPDRVSAQBALVHYFSALPSQLYQLPDESLFTVSGVRL 411
 DB 286 -----ESLFTVSGVRL 297
 QY 412 KPEMCDLASQYKGGHHPAQFSKCW 435
 DB 298 KPEMCDLASQYKGGHHPAQFSKCW 321

RESULT 9

ID ADM16427 standard; protein; 345 AA.
 AC ADM16427;

DT 17-JUN-2004 (first entry)

DE Human kinase protein, seq id 6.

XX Cytostatic; cardiact; neuroprotective; nootropic; antiinfertility;
 XX vulnery; antidiabetic; kinase; cancer; heart disease;
 XX Alzheimer's disease; infertility; wound; diabetes; neurological disease.

XX Homo sapiens.

XX WO2004024913-A1.

XX 25-MAR-2004.

XX 10-SEP-2003; 2003WO-JP011552.

XX 10-SEP-2002; 2002JP-00264345.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX (ZOG-) ZOGGENE CORP.

XX Kondo J, Kawai K, Miyama N, Nakajima M, Isogai T, Sugiyama T;
 PI Wakamatsu A, Irie R, Ishii S;
 DR MPI; 2004-270042/25.
 DR N-PDB; ADM16422.

XX Proteins of human origin having kinase activity, useful for prevention
 PT and treatment of kinase-associated diseases including cancer, heart
 PT disease and Alzheimer's disease.

PS Claim 1; SEQ ID NO 6; 105pp; Japanese.

XX The invention relates to four proteins of human origin (1) having kinase
 CC activity, and to proteins derived from these by addition, deletion and/or
 CC substitution of one or more amino acid residues, and having similar
 CC activity. The proteins and other aspects of the invention are useful for
 CC the prevention and treatment of kinase-associated diseases including
 CC cancer, heart disease, Alzheimer's disease, infertility, wounds, diabetes
 CC and neurological diseases. The current sequence represents a human
 CC protein having kinase activity.

XX Sequence 345 AA;

Query Match 70.8%; Score 1641.5; DB 8; Length 345;
 Best Local Similarity 93.8%; Pred. No. 2.6e-155;
 Matches 315; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

QY 1 MGOELCAKTVQRCSCYHCEGSEGAHSCRRSQPETTEAFKLTDLKEASCMTSPHPRGL 60
 DB 1 MGOELCAKTVQRCSCYHCEGSEGAHSCRRSQPETTEAFKLTDLKEASCMTSPHPRGL 60
 QY 61 QARARQKFKSKPRNSDCFOEDLRGCFQWRKSLPFGAASLYLNLEKLGESYATVYKG 120
 DB 61 QARARQKFKSKPRNSDCFOEDLRGCFQWRKSLPFGAASLYLNLEKLGESYATVYKG 120
 QY 121 ISIRINGQLVALKVIISMAEGVPPTAIRASLLKGLKANIYVLLHDIHTKTLTFVFEY 180
 DB 121 ISIRINGQLVALKVIISMAEGVPPTAIRASLLKGLKANIYVLLHDIHTKTLTFVFEY 180
 QY 181 NHTDLAQYMSQHFGGLHPHNVRLFMQLRLGLAYIHQHVLRDLKPNLLISHLGELK 240
 DB 181 N-----LFMQLRLGLAYIHQHVLRDLKPNLLISHLGELK 219
 QY 241 ADFGLARAKISPSQTSSEVYTLWYRPPDALLGATEYSELDIWGACIFIMFOGQPLF 300
 DB 220 ADFGLARAKISPSQTSSEVYTLWYRPPDALLGATEYSELDIWGACIFIMFOGQPLF 279
 QY 301 PGVSNILEQLEKIMEVLTGPTEDTWPVSKLPNNP 336
 DB 280 PGVSNILEQLEKIMEVLTGPTEDTWPVSKLPNNP 315

RESULT 10

ID ABP96082 standard; protein; 296 AA.

AC ABP96082;

DT 07-MAY-2003 (first entry)

DE Human protein kinase SEQ ID NO:70.

XX Human; protein kinase; enzyme; antiaesthetic; antiinflammatory;
 XX antidiabetic; antiparkinsonian; antigravine; cardiant; cytostatic;
 XX immunosuppressive; vulnerary; gene therapy; COPD; asthma; migraine;
 XX chronic obstructive pulmonary disease; non-insulin dependent diabetes;
 XX Parkinson's disease; myocardial infarction; inflammatory bowel disease;
 XX autoimmune disorder; allograft rejection; graft versus host disease;
 XX cancer; leukaemia; wound granulation.

XX Homo sapiens.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX (ZOG-) ZOGGENE CORP.

PN WO2003000901-A2.
 XX
 PD 03-JAN-2003.
 XX
 PF 24-JUN-2002; 2002WO-IB002358.
 XX
 XX 26-JUN-2001; 2001US-0301098P.
 PR 06-NOV-2001; 2001US-0332870P.
 XX
 PA (DECO-) DECODE GENETICS EHF.
 XX
 PI Martinez RM, Sigurdson GT;
 XX
 DR WPI; 2003-201429/19.
 DR N-PSDB; AB277160.
 XX
 PT New protein kinase genes and polypeptides, useful for diagnosing diseases
 PT associated with a protein kinase, or in gene therapy for treating e.g.
 PT Parkinson's disease, migraine, myocardial infarction, allograft rejection
 PT or cancers.
 PS
 PS Claim 9; Page 85; 258pp; English.
 XX
 CC AB277126 to AB277165 encode the human protein kinases given in ABP96048
 CC to ABP96087. The protein kinases have antiaesthetic, antiinflammatory,
 CC antidiabetic, antiparkinsonian, antimigraine, cardiant, cyostatic,
 CC immunosuppressive and vulnerary activities, and can be used in gene
 CC therapy. A protein kinase therapeutic agent from the present invention,
 CC particularly a protein kinase gene agonist or antagonist, can be used for
 CC treating a disease or condition associated with a protein kinase in an
 CC individual. These diseases include chronic obstructive pulmonary diseases
 CC (COPD), asthma, non-insulin dependent diabetes, Parkinson's disease,
 CC migraine, myocardial infarction, inflammatory bowel disease, autoimmune
 CC disorders (e.g. allograft rejection or graft vs. host disease), cancers
 CC (e.g. leukaemia) or wound granulation
 CC
 XX Sequence 296 AA;
 SQ
 Query Match 61.7%; Score 1432.5; DB 6; Length 296;
 Best Local Similarity 93.7%; Pred. No. 1.9e-134;
 Matches 282; Conservative 5; Mismatches 7; Indels 7; Gaps 3;
 QY 100 ASSYINLEKLGEGSVATYKGISRINQQLVALKVISNNAEGVPTFKIRASLLKGLKHA 159
 DB 1 ASSYINLEKLGEGSVATYKGISRINQQLVALKVISNNAEGVPTFKIRASLLKGLKHA 60
 QY 160 NIVLHDIHTKRETLTFEYEMHDDLAQYMSOHGGLHPHVRLFMFQLRGLAYIHQH 219
 DB 61 NIVLHDIHTKRETLTFEYEMVSYLAQYMSOHGGLHPHVRLFMFQLRGLAYIHQH 120
 QY 220 VLHRLDKPQNTLLISHLGELKLDLFGARAKSIPSGTYSSEVVTLMYRPPDALLGATEYSS 279
 DB 121 VLHRLDKPQNTLLISHLGELKLDLFGARAKSIPSGTYSSEVVTLMYRPPDALLGATEYSS 180
 QY 280 ELDIYGACIPIEMNGOQPLFPQVSNILBOLEKIMVILGVPTEDTWPQVSKLPVYNPE-W 338
 DB 181 ELDIYGACIPIEMNGOQPLFPQVSNILBOLEKIMVILGVPTEDTWPQVSKLPVYNPE-W 240
 QY 339 FFLPFRSLHVV-VNRLGRVPEAEPLASOMLKGPPRDVSAQELVHDFYFALPSQLYQL 397
 DB 241 -----RNSIFLSHFRSLGRVPEAEPLASOMLKGPPRDVSAQELVHDFYFALPSQLYQL 295
 QY 398 P 398
 DB 296 P 296
 RESULT 11
 AAB65641
 ID AAB65641 standard; protein; 266 AA.
 XX
 AC AAB65641;
 XX

DT 27-MAR-2001 (first entry)
 XX
 DB Novel protein kinase, SEQ ID NO: 168.
 XX
 KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiant; renal; antiinflammatory; antiaesthetic;
 KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200073469-A2.
 XX
 PD 07-DEC-2000.
 XX
 XX 26-MAY-2000; 2000WO-US014842.
 PF 28-MAY-1999; 99US-0136503P.
 PR
 XX
 PA (SUGEN-) SUGEN INC.
 XX
 PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
 XX
 DR WPI; 2001-032161/04.
 DR N-PSDB; AAF44668.
 XX
 PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers.
 PS
 PS Claim 10; Fig 1; 310pp; English.
 XX
 CC The present sequence is a novel protein kinase. The novel protein kinases
 CC and the nucleic acids that encode them may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomypathies, strokes, renal failure, oxidative-
 CC stress related disorder, chronic, inflammatory bowel disease, chronic
 CC inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,
 CC psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive
 CC disorders
 CC
 XX Sequence 266 AA;
 SQ
 Query Match 54.3%; Score 1260.5; DB 4; Length 266;
 Best Local Similarity 90.6%; Pred. No. 2.8e-117;
 Matches 241; Conservative 4; Mismatches 18; Indels 3; Gaps 1;
 QY 173 TLTFFVEYHNTDILAQYMSOHGGLHPHV--RLFMFQLRGLAYIHQHVLHRLDKPQN 229
 DB 1 TLTFFVEYHNTDILAEKASGNSNLMDSSQLMFMQLRGLAYIHQHVLHRLDKPQN 60
 QY 230 LILSHLGELKLDLFGARAKSIPSGTYSSEVVTLMYRPPDALLGATEYSSELDYGACI 289
 DB 61 LILSHLGELKLDLFGARAKSIPSGTYSSEVVTLMYRPPDALLGATEYSSELDYGACI 120
 QY 290 FIFMFOGQPLFPQVSNILBOLEKIMVILGVPTEDTWPQVSKLPVYNPEFPLPTRSLHV 349
 DB 121 FIFMFOGQPLFPQVSNILBOLEKIMVILGVPTEDTWPQVSKLPVYNPEFPLPTRSLHV 180
 QY 350 VNNRLGRVPEAEPLASOMLKGPPRDVSAQELVHDFYFALPSQLYQLPDESSLFTVSGV 409
 DB 181 VNNRLGRVPEAEPLASOMLKGPPRDVSAQELVHDFYFALPSQLYQLPDESSLFTVSGV 240

QY 410 RLKPEKCDLLASYOKGHHPAQPSKCM 435
 DB 241 RLKPEKCDLLASYOKGHHPAQPSKCM 266

RESULT 12
 ADI29248
 ID ADI29248 standard; protein; 266 AA.
 AC ADI29248;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human MARK3-associated protein #36.
 XX
 KW Human; antisense gene therapy; MARK3;
 KW MAP/microtubule affinity-regulating kinase 3; cancer;
 KW Alzheimer's disease; neurodegenerative disorder;
 KW hyperproliferative disorder; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN US2003232771-A1.
 PD 18-DEC-2003.
 XX
 PF 17-JUN-2002; 2002US-00174319.
 XX
 PR 17-JUN-2002; 2002US-00174319.
 XX
 PA (ISTS-) ISTS PHARM INC.
 XX
 PI Ward DT, Freier SM, Dobie KW;
 XX
 DR WPI: 2004-052188/05.
 DR N-PSDB; ADI29366.
 XX
 PT New antisense compound targeted to a nucleic acid molecule encoding
 PT microtubule-affinity-regulating kinases (MARK3), useful for modulating
 PT expression of MARK3 or for treating cancer or Alzheimer's disease.
 XX
 PS Disclosure; SEQ ID NO 160; 233bp; English.

CC The invention relates to a compound comprising a sequence comprising 8-80
 CC base pairs (bp) targeted to a nucleic acid encoding MARK3
 CC (MAP/microtubule affinity-regulating kinase 3), that specifically
 CC hybridises with the nucleic acid encoding MARK3 and inhibits expression
 CC of MARK3, i.e. is an antisense oligonucleotide (AO). Also included are a
 CC composition comprising the compound and a carrier or diluent, inhibiting
 CC the expression of MARK3 in cells or tissues, treating an animal having or
 CC suspected of having a disease or condition associated with MARK3 and
 CC screening for an antisense compound. The antisense oligonucleotide is
 CC useful for preparing a composition for treating hyperproliferative
 CC disorder, particularly cancer and neurodegenerative diseases e.g.
 CC Alzheimer's disease. The present sequence is a MARK3 associated protein
 CC included in the figures but not mentioned anywhere else in the
 CC specification.

SO Sequence 266 AA;

Query Match 54.3%; Score 1260.5; DB 8; Length 266;
 Best Local Similarity 90.6%; Pred. No. 2.8e-117;
 Matches 241; Conservative 4; Mismatches 18; Indels 3; Gaps 1;

QY 173 TLTFVEYHATDQAQMSQHFGHHPHNV---RLTFPQLRGALYHHQVHHRDKPON 229
 DB 1 TFTTFEYHATDQAQMSQHFGHHPHNV---RLTFPQLRGALYHHQVHHRDKPON 60

QY 230 LLTSHLGEKLADFGFLARAKSIPQTYSESVVTLWTRPPDALGATEYSSSLDIMGAGCT 289
 DB 61 LLTSHLGEKLADFGFLARAKSIPQTYSESVVTLWTRPPDALGATEYSSSLDIMGAGCT 120

QY 290 FIEMFOGOLFPQVSNILLEOLEKIMEVLGVPTEDTWPQSVSKLPNNPEWFPPLTPRSLHV 349
 DB 121 FIEMFOGOLFPQVSNILLEOLEKIMEVLGVPTEDTWPQSVSKLPNNPEWFPPLTPRSLHV 180

QY 350 VVNRRLGRVPEADLASQMLKGFPRDRVSAQELVHDYFSAIPSOYLDEESLFTVSGV 409
 DB 181 VVNRRLGRVPEADLASQMLKGFPRDRVSAQELVHDYFSAIPSOYLDEESLFTVSGV 240

QY 410 RLKPEKCDLLASYOKGHHPAQPSKCM 435
 DB 241 RLKPEKCDLLASYOKGHHPAQPSKCM 266

RESULT 13
 ABR57358
 ID ABR57358 standard; protein; 469 AA.
 AC ABR57358;
 XX
 DT 09-SEP-2003 (first entry)
 XX
 DE Mouse serine/threonine protein kinase SEQ ID NO:3.
 XX
 KW Mouse; serine/threonine protein kinase; kinase; enzyme; cytostatic;
 KW antidiabetic; neuroprotective; nootropic; antiparkinsonian; analgesic;
 KW cerebroprotective; antiinflammatory; gastrointestinal; hepatotropic;
 KW gynaecological; uropathic; dermatological; gene therapy; cancer;
 KW diabetes; central nervous system disorder; CNS disorder; liver disease;
 KW respiratory disorder; chronic obstructive pulmonary disease; stroke;
 KW cardiovascular disorder; dermatological disorder; urological disorder;
 KW gastrointestinal disease; haematological disorder; Alzheimer's disease;
 KW musculoskeletal disorder; reproductive disorder; Parkinson's disease;
 KW neuropathic pain.
 XX
 OS Mus musculus.
 XX
 PN WO2003046167-A1.
 PD 05-JUN-2003.
 XX
 PF 26-NOV-2002; 2002WO-EP013268.
 XX
 PR 27-NOV-2001; 2001US-0333131P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Koehler RH;
 XX
 DR WPI: 2003-505196/47.
 XX
 PT New polynucleotide encoding a serine/threonine protein kinase
 PT polypeptide, useful for diagnosing, preventing or treating diseases
 PT associated with serine/threonine protein kinase dysfunction, e.g. cancer
 PT or diabetes.

XX Disclosure; Fig 3; 196bp; English.

XX The present sequence represents a mouse serine/threonine protein kinase
 CC (1). (1) has cytostatic, antidiabetic, neuroprotective, antiparkinsonian,
 CC nootropic, cerebroprotective, analgesic, antiinflammatory, hepatotropic,
 CC gastrointestinal, gynaecological, uropathic and dermatological
 CC activities, and can be used in gene therapy. Serine/threonine protein
 CC kinase polynucleotide and polypeptide sequences can be used in
 CC diagnosing, preventing, ameliorating or treating diseases associated with
 CC serine/threonine protein kinase dysfunction. They may also be used to
 CC identify test compounds that may act, for example, as activators or
 CC inhibitors at the enzyme's active site. The human serine/threonine
 CC protein kinase and its fragments are also useful in raising specific
 CC antibodies that can block the enzyme and effectively reduce its activity.
 CC Human serine/threonine protein kinase sequences can be used in the
 CC preparation of a medicament for modulating the activity of a serine/
 CC threonine protein kinase in a disease, e.g. cancer, diabetes, a central
 CC nervous system (CNS) disorder, a respiratory disorder (including chronic

CC obstructive pulmonary disease), a cardiovascular disorder, a
 CC dermatological disorder, a gastrointestinal or liver disease, a
 CC haematological disorder, a musculoskeletal disorder, a reproductive
 CC disorder, or a urological disorder. CNS disorders may include Alzheimer's
 CC disease, Parkinson's disease, stroke or neuropathic pain
 XX

SO Sequence 469 AA;

Query Match 51.7%; Score 1198.5; DB 7; Length 469;
 Best Local Similarity 55.3%; Pred. No. 1.1e-110;
 Matches 246; Conservative 55; Mismatches 103; Indels 41; Gaps 5;

QY 23 GEANSCRSQPER-TEAATLTLDLKASCSMTSPHRLGQA-----ARA 65
 DB 14 GKMKKLRLTSSFSRLAKKEDTTFDEICVTMSTRNCGTDSVIKHLDTIPEDKKVRV 73
 QY 66 QKPKS-----KRPKS-NSDCFOEBDLRGQFQWRK-----SLPPGAAS 101
 DB 74 QRTOSTFPPEKPRANQVKRVHSENNACINFKSSAGKSPKVRHSSPSPSPKFGKAD 133
 QY 102 SYLNLEKLGEGSYATVYKGISRINGOLVALKVISMNAEGVPTAIRASLKGKXANI 161
 DB 134 SYEKLEKLGEGSYATVYKGISRINGOLVALKVIRLQEEGPTAIRASLKGKXANI 193
 QY 162 VLLHDIHHTKETLTFVEYMTDLAQYMSOHFGSLPHNVRLEMFQLRGLAYIHQHYL 221
 DB 194 VLLHDIHHTKETLTFVEYMTDLAQYMSOHFGSLPHNVRLEMFQLRGLAYIHQHYL 253
 QY 222 HRDLKQNTLLISHGELKADFGALARAKSIPQYSSSEVTLWYRPPDALLGATEYSSSL 281
 DB 254 HRDLKQNTLLISHGELKADFGALARAKSIPSHSYSEVTLWYRPPDALLGATEYSSSL 313
 QY 282 DWAGACIFIEWFOGQPLPGVSNILBOLEKIWEVTCVPTEDTWPGVSKLPYNPEWPL 341
 DB 314 DWAGACIFIEWFOGQPLPGVSNILBOLEKIWEVTCVPTEDTWPGVSKLPYNPEWPL 373
 QY 342 PTPRSILHVWNRLLGRVPEAEADLASOMLKGPBRDVSQAQELVHDYSALPSQLYOLPDEE 401
 DB 374 YSSKSLRQAWNKLSYVNHAEADLASKLQCSFKNRLSAQALSHYFSDLPRLMELTDMS 433
 QY 402 SLFTVSGVRLKPEMCDLASQYQKH 426
 DB 434 SIFTVPNVRLLQPEAGESMRAFGKNN 458

RESULT 14

ADJ76287
 ID ADJ76287 standard; protein; 469 AA.

AC ADJ76287;

DT 20-MAY-2004 (first entry)

DE Marker gene related amino acid sequence SEQ ID NO:1539.

XX bronchial asthma; chronic obstructive pulmonary disease;
 KM respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
 KW gene therapy; marker.

OS Mus musculus.

PN EP1394274-A2.

PD 03-MAR-2004.

PF 04-AUG-2003; 2003BP-00254857.

PR 06-AUG-2002; 2002JP-00229312.

PR 20-MAR-2003; 2003JP-00077212.

XX (GENO-) GENOX RES INC.

XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhabara K;

XX MPI, 2004-193155/19.

DR Testing for bronchial asthma or chronic obstructive pulmonary disease by
 XX comparing the expression level of a marker gene in a biological sample
 PT from a subject with the expression level of the gene in a sample from a
 PT healthy subject.

PS Example 11; SEQ ID NO 1539; 241p; English.

XX The present invention describes a method of testing for bronchial asthma
 XX or chronic obstructive pulmonary disease. The method comprises
 CC determining the expression level of a marker gene in a biological sample
 CC from a subject, comparing the expression level determined with the
 CC expression level of the marker gene in a biological sample from a healthy
 CC subject, and judging whether the subject has bronchial asthma or chronic
 CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
 CC genes (81) whose expression levels increase when respiratory epithelial
 CC cells are stimulated with interleukin-13; or (b) a group of genes (82)
 CC whose expression levels decrease when respiratory epithelial cells are
 CC stimulated with interleukin-13. Also described: (1) a reagent (1) for
 CC testing for bronchial asthma or chronic obstructive pulmonary disease;
 CC (2) a kit for screening for a candidate compound for a therapeutic agent
 CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
 CC an animal model for bronchial asthma or chronic obstructive pulmonary
 CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
 CC method for producing an animal model for bronchial asthma or chronic
 CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
 CC asthma or chronic obstructive pulmonary disease, comprising the compound,
 CC a marker gene or an antisense nucleic acid corresponding to a portion of
 CC the marker gene, a ribozyme, a polynucleotide that suppresses the
 CC expression of the gene through an RNAi effect or an antibody recognising
 CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
 CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
 CC probe has been immobilised to assay a marker gene. (1) has respiratory
 CC and antiasthmatic activities, and can be used in gene therapy. The method
 CC is useful for testing for or screening for a therapeutic agent for
 CC bronchial asthma or chronic obstructive pulmonary disease. The present
 CC sequence is used in the exemplification of the present invention.

XX SO Sequence 469 AA;

Query Match 51.7%; Score 1198.5; DB 8; Length 469;
 Best Local Similarity 55.3%; Pred. No. 1.1e-110;
 Matches 246; Conservative 55; Mismatches 103; Indels 41; Gaps 5;

QY 23 GEANSCRSQPER-TEAATLTLDLKASCSMTSPHRLGQA-----ARA 65
 DB 14 GKMKKLRLTSSFSRLAKKEDTTFDEICVTMSTRNCGTDSVIKHLDTIPEDKKVRV 73
 QY 66 QKPKS-----KRPKS-NSDCFOEBDLRGQFQWRK-----SLPPGAAS 101
 DB 74 QRTOSTFPPEKPRANQVKRVHSENNACINFKSSAGKSPKVRHSSPSPSPKFGKAD 133
 QY 102 SYLNLEKLGEGSYATVYKGISRINGOLVALKVISMNAEGVPTAIRASLKGKXANI 161
 DB 134 SYEKLEKLGEGSYATVYKGISRINGOLVALKVIRLQEEGPTAIRASLKGKXANI 193
 QY 162 VLLHDIHHTKETLTFVEYMTDLAQYMSOHFGSLPHNVRLEMFQLRGLAYIHQHYL 221
 DB 194 VLLHDIHHTKETLTFVEYMTDLAQYMSOHFGSLPHNVRLEMFQLRGLAYIHQHYL 253
 QY 222 HRDLKQNTLLISHGELKADFGALARAKSIPQYSSSEVTLWYRPPDALLGATEYSSSL 281
 DB 254 HRDLKQNTLLISHGELKADFGALARAKSIPSHSYSEVTLWYRPPDALLGATEYSSSL 313
 QY 282 DWAGACIFIEWFOGQPLPGVSNILBOLEKIWEVTCVPTEDTWPGVSKLPYNPEWPL 341
 DB 314 DWAGACIFIEWFOGQPLPGVSNILBOLEKIWEVTCVPTEDTWPGVSKLPYNPEWPL 373
 QY 342 PTPRSILHVWNRLLGRVPEAEADLASOMLKGPBRDVSQAQELVHDYSALPSQLYOLPDEE 401
 DB 374 YSSKSLRQAWNKLSYVNHAEADLASKLQCSFKNRLSAQALSHYFSDLPRLMELTDMS 433

OY 402 SLFTVSGVRLKPEKCDLLASYOKH 426
 DB 434 SIFTVPNVRLQPEAGESMRAFGKN 458

RESULT 15

AA018617 standard; protein; 330 AA.

AC AAO18617;

DT 24-OCT-2002 (first entry)

DE Human pftaire family kinase related murine protein-SEQ ID NO: 8.

KW Human; pftaire family kinase; kinase; enzyme; testis; brain; cytotatic; uterus endometrium adenocarcinoma; lung fibroblast; splice form;

KW kidney renal cell adenocarcinoma; gene therapy; mouse.

OS Mus musculus.

PN W0200261060-A2.

PD 08-AUG-2002.

PF 17-JAN-2002; 2002WO-US001106.

PR 31-JAN-2001; 2001US-0265151P.

PR 09-MAR-2001; 2001US-00801861.

PA (PEKE) PE CORP NY.

PI Yan C, Ketchum K, Di Francesco V, Beasley EM;

DR WPI; 2002-608515/65.

PT New human kinase peptide and nucleic acid molecule, useful for treating disorders associated with abnormal expression of kinase protein, e.g.

PT adenocarcinoma of uterus or lung, in drug screening assays and pharmacogenomic analysis.

PT pharmacogenomic analysis.

PS Disclosure; Page 129-130; 131pp; English.

CC The present invention provides the protein, cDNA and gene sequences of two splice variants of a human pftaire family kinase. The sequences are specifically expressed in the human testis, brain, uterus endometrium adenocarcinomas, lung fibroblasts, kidney renal cell adenocarcinomas, and can be used to treat related diseases. The present sequence is a murine protein shown in the invention. Note: The present sequence is not described further in the specification

CC Sequence 330 AA;

Query Match 51.5%; Score 1195; DB 5; Length 330;

Best Local Similarity 68.2%; Pred. No. 1.4e-110; Mismatches 60; Indels 0; Gaps 0;

DB 97 FGAAASYLULKLEKSGSVATYVKGISRINGOLVALKVISMNAEEGVPTAIRBASLLKGL 156
 1 FGKAOSYEKLEKSGSVATYVKGISKVGKVALKVIROEBEGTPTAIRBASLLKGL 60

OY 157 KHANIVLHDIHTKETLTTFVEYMTDLAQYMSQHPGGLHPHNVLFMFOLLRGLAYIH 216
 DB 61 KHANIVLHDIHTKETLTTFVEYMTDLAQYMSQHPGGLHPHNVLFMFOLLRGLAYIH 120

OY 217 HQHVLHRLDKPQNLISHLGEKLADFGJARAKSIPQYSSSVTLWYRPPDALLGATE 276
 DB 121 ORYIHLHRLDKPQNLISHLGEKLADFGJARAKSIPQYSSSVTLWYRPPDALLGATE 180

OY 277 YSSELIDWAGCIFIEMFGQPLFPGVSNILBQLEKIWEVLGVPTEDTPGVSCLPNYP 336
 DB 181 YSTCLDMWAGCIFIEMFGQVLAAPGMKDIQDOLERIFLVLTGTPNEDTPGVSCLPHKXP 240

OY 337 EWFPLPTRSLSHVNNRRLGRVPEADLASOMLKGFPRDRVSNQOALVHDFSAIPLSOLXQ 396
 DB 241 ERFTYSSKSLRQAVNKLSYVNHADLASKLQCSPPKRLSAQALSHYFSDLPRLWE 300
 OY 397 LPDESLFTVSGVRLKPEKCDLLASYOKH 426
 DB 301 LTDMSIFTVPNVRLQPEAGESMRAFGKN 330

Search completed: December 27, 2004, 12:16:10
 Job time : 76 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2004, 17:48:22 ; Search time 421.082 Seconds
(without alignments)
17296.376 Million cell updates/sec

Title: US-10-786-065-4

Perfect score: 1308
Sequence: 1 atcgagtcgaagcgcgtgctgc.....agttacgaatcgtcgtgta 1308

Scoring table:
Gapop 10.0, Gapext 1.0

Searched: 4105333 seqs, 2784095677 residues

Total number of hits satisfying chosen parameters: 8210666

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1308	100.0	1308	9	US-09-801-861-4
2	1308	100.0	1308	14	US-10-224-562-4
3	1308	100.0	2250	18	US-10-618-941-20
4	1009	77.1	2203	9	US-09-801-861-1
5	1009	77.1	2203	14	US-10-224-562-1
6	887.6	67.9	1376	16	US-10-466-759-15
7	851.8	65.1	3210	16	US-10-425-114-26244
8	481	36.8	481	9	US-09-948-802-14
9	481	36.8	481	15	US-10-121-925-14
10	438	33.5	2467	15	US-10-153-242-3
11	437.8	33.5	1372	15	US-10-153-242-1
12	436.2	33.3	1549	15	US-10-153-242-13

13	436.2	33.3	4957	18	US-10-723-860-406	Sequence 406, App
14	436.2	33.3	5161	18	US-10-723-860-5096	Sequence 5096, App
15	299.8	22.9	3024	15	US-10-296-115-483	Sequence 483, App
16	297.8	22.8	1745	15	US-10-017-621-3	Sequence 3, App1
17	297.8	22.8	1745	15	US-10-354-358-63	Sequence 63, App1
18	297.8	22.8	1745	15	US-10-172-118-1254	Sequence 1254, App
19	297.8	22.8	1745	16	US-10-342-984-1254	Sequence 1254, App
20	297.8	22.8	2382	15	US-10-112-944-42	Sequence 42, App
21	297.8	22.8	2987	15	US-10-017-621-89	Sequence 89, App1
22	297.8	22.8	3131	18	US-10-723-860-2262	Sequence 2262, App
23	297.8	22.8	3182	15	US-10-017-621-88	Sequence 88, App1
24	297.4	22.7	2825	9	US-09-925-298-72	Sequence 72, App1
25	297.4	22.7	2825	13	US-10-102-806-72	Sequence 324, App
26	296.6	22.7	2580	13	US-10-098-841-324	Sequence 31, App1
27	296.6	22.7	2858	16	US-10-258-106-31	Sequence 31, App1
28	296.6	22.7	3094	18	US-10-851-921-7	Sequence 7, App1
29	296.6	22.7	3229	18	US-10-618-941-19	Sequence 19, App1
30	296.6	22.7	3243	13	US-10-087-192-1301	Sequence 1301, App
31	296.2	22.6	2868	15	US-10-037-270-53	Sequence 53, App1
32	296.2	22.6	2868	15	US-10-117-722-53	Sequence 53, App1
33	291.8	22.3	1125	15	US-10-325-430-5	Sequence 5, App1
34	291.8	22.3	1242	15	US-10-325-430-4	Sequence 4, App1
35	291.8	22.3	1242	16	US-10-305-720-1283	Sequence 1283, App
36	291.8	22.3	1242	18	US-10-757-262-49	Sequence 49, App1
37	290.8	22.2	2653	13	US-10-087-192-1298	Sequence 1298, App
38	285.2	21.8	3605	18	US-10-723-860-6480	Sequence 6480, App
39	269.2	20.6	1738	15	US-10-354-358-85	Sequence 85, App1
40	269.2	20.6	1738	15	US-10-160-787-4	Sequence 4, App1
41	244	18.7	2785	13	US-10-087-192-1505	Sequence 1505, App
42	231.2	17.7	667	9	US-09-764-868-183	Sequence 183, App
43	231.2	17.7	667	11	US-09-764-875-195	Sequence 395, App
44	227	17.4	1287	11	US-09-764-875-78	Sequence 78, App
45	221.4	16.9	918	9	US-09-801-368-261	Sequence 261, App

ALIGNMENTS

US-09-801-861-4	Application US/09801861
Sequence 4, Appl1	
Patent No. US20020119544A1	
GENERAL INFORMATION:	
APPLICANT: YAN, Chunhua et al.	
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC	
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES	
TITLE OF INVENTION: THEREOF	
FILE REFERENCE: CLO01098	
CURRENT APPLICATION NUMBER: US/09/801,861	
CURRENT FILING DATE: 2001-03-09	
NUMBER OF SEQ ID NOS: 10	
SOFTWARE: FastSeq for Windows Version 4.0	
SEQ ID NO 4	
LENGTH: 1308	
TYPE: DNA	
ORGANISM: Human	
US-09-801-861-4	
Query Match	100.0%; Score 1308; DB 9; Length 1308;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGGGTCAGAGCTGTGTCAGAAAGCTGTACAGCTGTGATGAGCTGTCTACCATTTGCA 60
DB	1 ATGGGTCAGAGCTGTGTCAGAAAGCTGTACAGCTGTGATGAGCTGTCTACCATTTGCA 60
QY	61 GAGGAGGCGGAGGACACACAGCTGTGAGAGGATCCAGCTGAGACCGAGGAGGCTGCTTC 120
DB	61 GAGGAGGCGGAGGACACACAGCTGTGAGAGGATCCAGCTGAGACCGAGGAGGCTGCTTC 120
QY	121 AAGCTTACAGAGCTTAAAGAGCATCATGTTCCATGACTTTCATCCACCCAGGGGACTT 180
DB	121 AAGCTTACAGAGCTTAAAGAGCATCATGTTCCATGACTTTCATCCACCCAGGGGACTT 180

QY 181 CAAGTCCCCGTCGCAAGTTCAGAGTAAAGGCCAGAGTACAGTATGTTT 240
DB 181 CAAGTCCCCGTCGCAAGTTCAGAGTAAAGGCCAGAGTACAGTATGTTT 240
QY 241 CAGGAAGAGATCTGAGGAGGGTTTTCAGTGGAGGAGAGCTCCCTTTTGGGAGGCC 300
DB 241 CAGGAAGAGATCTGAGGAGGGTTTTCAGTGGAGGAGAGCTCCCTTTTGGGAGGCC 300
QY 301 TCATCTTACTTGAACCTTGAAGAGCTGGGTGAAGGCTTTATGCGACAGTTTCAAGGGG 360
DB 301 TCATCTTACTTGAACCTTGAAGAGCTGGGTGAAGGCTTTATGCGACAGTTTCAAGGGG 360
QY 361 ATTAGAGAAATAATGAGCAACTAGTGGCTTTAAAGTATCGAGTGAATGAGAGAA 420
DB 361 ATTAGAGAAATAATGAGCAACTAGTGGCTTTAAAGTATCGAGTGAATGAGAGAA 420
QY 421 GGAGTCCCATTTACAGCTATCCGAGAGGCTTCTCTGAGAGGGTTGAAACATGCCAAT 480
DB 421 GGAGTCCCATTTACAGCTATCCGAGAGGCTTCTCTGAGAGGGTTGAAACATGCCAAT 480
QY 481 ATTGTCCTCTGATGACATATCCACCAAGAGACATGACATTCGTTTTGAATAC 540
DB 481 ATTGTCCTCTGATGACATATCCACCAAGAGACATGACATTCGTTTTGAATAC 540
QY 541 ATGCAACAAGACCTGGGCCAGATATATGTCCTGAGAGGGCTTCACTCTCAAT 600
DB 541 ATGCAACAAGACCTGGGCCAGATATATGTCCTGAGAGGGCTTCACTCTCAAT 600
QY 601 GTGAGACTTTTCACTTCACTTTTGCAGGGGCTGAGCTGATCCACCAACAACGTT 660
DB 601 GTGAGACTTTTCACTTCACTTTTGCAGGGGCTGAGCTGATCCACCAACAACGTT 660
QY 661 CTTCACAGGAGCTTGAACTCAGAACTTATCTCATGATCACTGGAGAGCTCAAACTG 720
DB 661 CTTCACAGGAGCTTGAACTCAGAACTTATCTCATGATCACTGGAGAGCTCAAACTG 720
QY 721 GCGATTTTGTCTTGGCCGGGCGCAAGTCACTCCAGAGGAGCACTCTTCAGAAAGTC 780
DB 721 GCGATTTTGTCTTGGCCGGGCGCAAGTCACTCCAGAGGAGCACTCTTCAGAAAGTC 780
QY 781 GTGACCCCTGTGATCCGGGCCCCCTGATGCTTGTCTGGAGGCCACTGAATATTCCTGAG 840
DB 781 GTGACCCCTGTGATCCGGGCCCCCTGATGCTTGTCTGGAGGCCACTGAATATTCCTGAG 840
QY 841 CTGACATATGGGGTGCAGGCTGATCTTTATTGAATGTTCCAGGGTCAACTTTGTTT 900
DB 841 CTGACATATGGGGTGCAGGCTGATCTTTATTGAATGTTCCAGGGTCAACTTTGTTT 900
QY 901 CCGGGGTTTCCAAACATCTTGAACAGCTGGAAGAAATCTGGGAGGTCGGGAGTCCCT 960
DB 901 CCGGGGTTTCCAAACATCTTGAACAGCTGGAAGAAATCTGGGAGGTCGGGAGTCCCT 960
QY 961 ACAGAGGATCTTGGCCGGGAGTCTCCAAAGCTACTTACATCAATCAGAAATGGTCCCA 1020
DB 961 ACAGAGGATCTTGGCCGGGAGTCTCCAAAGCTACTTACATCAATCAGAAATGGTCCCA 1020
QY 1021 CTGCTTACGCTTGAAGCTTCACTGTTGTGGAACAGGCTGGGCAAGGTTCTGAAAGCT 1080
DB 1021 CTGCTTACGCTTGAAGCTTCACTGTTGTGGAACAGGCTGGGCAAGGTTCTGAAAGCT 1080
QY 1081 GAAAGACTGGGCTCCCAATGCTTAAAGGCTTTCCAGAGACCGGCTCTCCGCGCCAGGAA 1140
DB 1081 GAAAGACTGGGCTCCCAATGCTTAAAGGCTTTCCAGAGACCGGCTCTCCGCGCCAGGAA 1140
QY 1141 GCACTTGTTCATGATATTTACAGGCTGCGCATCTCAGCTGTACAGCTTCTCTGATGAG 1200
DB 1141 GCACTTGTTCATGATATTTACAGGCTGCGCATCTCAGCTGTACAGCTTCTCTGATGAG 1200
QY 1201 GAGTCTTGTTCATGATTCAGAGTGAAGGCTTAAGCCAGAAATGTGTGACTTTTGGCC 1260
DB 1201 GAGTCTTGTTCATGATTCAGAGTGAAGGCTTAAGCCAGAAATGTGTGACTTTTGGCC 1260

QY 1261 TCCTACAGGAAGGTCAACCCAGCCAGCTTTAGCAATGCTGGTGA 1308
DB 1261 TCCTACAGGAAGGTCAACCCAGCCAGCTTTAGCAATGCTGGTGA 1308

RESULT 2
US-10-224-562-4
Sequence 4, Application US/10224562
Publication No. US20030022229A1
GENERAL INFORMATION:
APPLICANT: VAN, Chunhua et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: CL001098D1V
CURRENT APPLICATION NUMBER: US/10/224,562
NUMBER OF FILING DATE: 2002-08-21
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 1308
TYPE: DNA
ORGANISM: Human
US-10-224-562-4

Query Match 100.0%; Score 1308; DB 14; Length 1308;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTCAAGAGCTGTGTCAGAAAGCTGATCAGCTGTGATGAGCTGTACCATTTGTTCA 60
DB 1 ATGGGTCAAGAGCTGTGTCAGAAAGCTGATCAGCTGTGATGAGCTGTACCATTTGTTCA 60
QY 61 GAGGAGGCGGAGCAACAGCTGTGAGAGAGTCAAGCTGTGAGACCAAGGAGCTGCGTTT 120
DB 61 GAGGAGGCGGAGCAACAGCTGTGAGAGAGTCAAGCTGTGAGAGCAACGAGGAGCTGCGTTT 120
QY 121 AAGCTTAACAGACTTAAAGAAAGCAATCATGTTCCATGACTTCACTTCAACCCAGGGGACT 180
DB 121 AAGCTTAACAGACTTAAAGAAAGCAATCATGTTCCATGACTTCACTTCAACCCAGGGGACT 180
QY 181 CAAGTCCCGTGGCCAGAGTTCAGAGTAAAGGCCACGAGTAAACAGTATGTTT 240
DB 181 CAAGTCCCGTGGCCAGAGTTCAGAGTAAAGGCCACGAGTAAACAGTATGTTT 240
QY 241 CAGGAAGAGATCTGAGGAGGGTTTCACTGAGAGAAAGCTTCCCTTTGGGAGGCC 300
DB 241 CAGGAAGAGATCTGAGGAGGGTTTCACTGAGAGAAAGCTTCCCTTTGGGAGGCC 300
QY 301 TCATCTTACTTGAACCTTGAAGAGCTGGGTGAAGGCTTTATGCGACAGTTTCAAGGGG 360
DB 301 TCATCTTACTTGAACCTTGAAGAGCTGGGTGAAGGCTTTATGCGACAGTTTCAAGGGG 360
QY 361 ATTAGAGAAATAATGAGCAACTAGTGGCTTTAAAGTATCGAGTGAATGAGAGAA 420
DB 361 ATTAGAGAAATAATGAGCAACTAGTGGCTTTAAAGTATCGAGTGAATGAGAGAA 420
QY 421 GGAGTCCCATTTACAGCTATCCGAGAGGCTTCTCTGAGAGGGTTGAAACATGCCAAT 480
DB 421 GGAGTCCCATTTACAGCTATCCGAGAGGCTTCTCTGAGAGGGTTGAAACATGCCAAT 480
QY 481 ATTGTCCTCTGATGACATATCCACCAAGAGACATGACATTCGTTTTGAATAC 540
DB 481 ATTGTCCTCTGATGACATATCCACCAAGAGACATGACATTCGTTTTGAATAC 540
QY 541 ATGCAACAAGACCTGGGCCAGATATATGTCCTGAGAGGGCTTCACTCTCAAT 600
DB 541 ATGCAACAAGACCTGGGCCAGATATATGTCCTGAGAGGGCTTCACTCTCAAT 600
QY 601 GTGAGACTTTTCACTTCACTTTTGCAGGGGCTGAGCTGATCCACCAACAACGTT 660
DB 601 GTGAGACTTTTCACTTCACTTTTGCAGGGGCTGAGCTGATCCACCAACAACGTT 660

QY 661 CTTCAAGGAGCTGAAACCTCAGAACTTACTCATGTCACCTGGAGAGCTCAAACTG 720
DB 661 CTTCAAGGAGCTGAAACCTCAGAACTTACTCATGTCACCTGGAGAGCTCAAACTG 720
QY 721 GCTGATTTTGGTCTTGCCCGGGCAAGTCCATTCCAGCAGACATATCTTCAGAAATC 780
DB 721 GCTGATTTTGGTCTTGCCCGGGCAAGTCCATTCCAGCAGACATATCTTCAGAAATC 780
QY 781 GTGACCTCTGTGTACCGGCCCCCTGATGCTTGTCTGGAGCCATGATATTTCTGTAG 840
DB 781 GTGACCTCTGTGTACCGGCCCCCTGATGCTTGTCTGGAGCCATGATATTTCTGTAG 840
QY 841 CTGACATATGGGGGTGACAGCTGATCTTTATTTGAAATGTCACAGGGTCAACCTTTGTT 900
DB 841 CTGACATATGGGGGTGACAGCTGATCTTTATTTGAAATGTCACAGGGTCAACCTTTGTT 900
QY 901 CTTGGGGTTTCCAACTCCTTGAAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCT 960
DB 901 CTTGGGGTTTCCAACTCCTTGAAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCT 960
QY 961 ACAGAGATCTTGGCCGGAGTCTTCAGAGCTAACCTAACATCAGATGTTCCCA 1020
DB 961 ACAGAGATCTTGGCCGGAGTCTTCAGAGCTAACCTAACATCAGATGTTCCCA 1020
QY 1021 CTGCTACGCTCGAAGCTTGTCTGGAACAGGCTGGGCAAGGTTCTGAAAGT 1080
DB 1021 CTGCTACGCTCGAAGCTTGTCTGGAACAGGCTGGGCAAGGTTCTGAAAGT 1080
QY 1081 GAAAGCTGGCTCCAGATGCTAAAGGCTTCCAGAGACCGCTCTCCGCCAGAA 1140
DB 1081 GAAAGCTGGCTCCAGATGCTAAAGGCTTCCAGAGACCGCTCTCCGCCAGAA 1140
QY 1141 GCACTTGTTCATGATTTATTTACGCGCCCTGCCATCTCAGCTGTACCACTTCTGTAG 1200
DB 1141 GCACTTGTTCATGATTTATTTACGCGCCCTGCCATCTCAGCTGTACCACTTCTGTAG 1200
QY 1201 GAGCTTGTTCATGATTTACAGAGTGAAGCTAAAGCCAGAAATGTTGACCTTTGGCC 1260
DB 1201 GAGCTTGTTCATGATTTACAGAGTGAAGCTAAAGCCAGAAATGTTGACCTTTGGCC 1260
QY 1261 TCTTACAGAAAGTCAACCAAGCCAGCTTACCAATGCTGTGA 1308
DB 1261 TCTTACAGAAAGTCAACCAAGCCAGCTTACCAATGCTGTGA 1308

RESULT 3
US-10-618-941-20
; Sequence 20, Application US/10618941
; Publication No. US2004019792A1
; GENERAL INFORMATION:
; APPLICANT: WHITE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 2250
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-618-941-20

Query Match 100.0%; Score 1308; DB 18; Length 2250;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTCAGAGCTGTGTCAAAAGTGTACAGCTGTAGTGAAGCTTACCATTTGTCA 60
DB 1 ATGGGTCAGAGCTGTGTCAAAAGTGTACAGCTGTAGTGAAGCTTACCATTTGTCA 60

DB 45 ATGGGTCAGAGCTGTGTCAAAAGTGTACAGCTGTAGTGAAGCTTACCATTTGTCA 104
QY 61 GGGGAGGCGAGGCAACAGCTGTGGAGAGTCAAGCTTGAAACCAAGGAGCTGGTTC 120
DB 105 GAGGAGGGGAGGCAACAGCTGTGGAGAGTCAAGCTTGAAACCAAGGAGCTGGTTC 164
QY 121 AAGCTAACAGACTTAAAGATCATGTTCCATGACTTCACTTCAACCCAGGGACTT 180
DB 165 AAGCTAACAGACTTAAAGATCATGTTCCATGACTTCACTTCAACCCAGGGACTT 224
QY 181 CAGGTGCCCCGTCAGAGTTCAGAGTAAAGGCCACGAGTAAAGTATTTGTTT 240
DB 225 CAGGTGCCCCGTCAGAGTTCAGAGTAAAGGCCACGAGTAAAGTATTTGTTT 284
QY 241 CAGGAAAGAGATCTGAGCGAGGTTTTCAGTGAAGAAAGGCTTCCCTTTGGGCAAGC 300
DB 285 CAGGAAAGAGATCTGAGCGAGGTTTTCAGTGAAGAAAGGCTTCCCTTTGGGCAAGC 344
QY 301 TCATCTTACTTGAACCTTGAGAAAGCTGGAGTGAAGGCTCTTATGCGACAGTTTACAAGGG 360
DB 345 TCATCTTACTTGAACCTTGAGAAAGCTGGAGTGAAGGCTCTTATGCGACAGTTTACAAGGG 404
QY 361 ATTACAGAAATTAATGAACTAAGTGTCTTAAAGTATCAGCATGATGATGAGAGAA 420
DB 405 ATTACAGAAATTAATGAACTAAGTGTCTTAAAGTATCAGCATGATGATGAGAGAA 464
QY 421 GAGTCCCATTTACAGCTATCCGAGAGCTTCTCTGAGAGGTTTGAACATGCGAAT 480
DB 465 GAGTCCCATTTACAGCTATCCGAGAGCTTCTCTGAGAGGTTTGAACATGCGAAT 524
QY 481 ATTGCTCTCGAGACATTAATCCACCAAGAGAGACATGACATTTGTTTGAATAC 540
DB 525 ATTGCTCTCGAGACATTAATCCACCAAGAGAGACATGACATTTGTTTGAATAC 584
QY 541 ATGCAACAGACCTGCGCCAGTATATGCTCAGCATCCAGAGGAGCTTCACTAAT 600
DB 585 ATGCAACAGACCTGCGCCAGTATATGCTCAGCATCCAGAGGAGCTTCACTAAT 644
QY 601 GTCAAGCTTTTATGTTTCACTTTTGGGGGCTGGGCTGATCCACCAACACGTT 660
DB 645 GTCAAGCTTTTATGTTTCACTTTTGGGGGCTGGGCTGATCCACCAACACGTT 704
QY 661 CTTCAAGGAGCTGAAACCTCAGAACTTACTCATGTCACCTGGAGAGCTCAAACTG 720
DB 705 CTTCAAGGAGCTGAAACCTCAGAACTTACTCATGTCACCTGGAGAGCTCAAACTG 764
QY 721 GCTGATTTTGGTCTTGCCCGGGCAAGTCCATTCCAGCAGACATATCTTCAGAAATC 780
DB 765 GCTGATTTTGGTCTTGCCCGGGCAAGTCCATTCCAGCAGACATATCTTCAGAAATC 824
QY 781 GTGACCTCTGTGTACCGGCCCCCTGATGCTTGTCTGGAGCCATGATATTTCTGTAG 840
DB 825 GTGACCTCTGTGTACCGGCCCCCTGATGCTTGTCTGGAGCCATGATATTTCTGTAG 884
QY 841 CTGACATATGGGGGTGACAGCTGATCTTTATTTGAAATGTCACAGGGTCAACCTTTGTT 900
DB 885 CTGACATATGGGGGTGACAGCTGATCTTTATTTGAAATGTCACAGGGTCAACCTTTGTT 944
QY 901 CTTGGGGTTTCCAACTCCTTGAAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCT 960
DB 945 CTTGGGGTTTCCAACTCCTTGAAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCT 1004
QY 961 ACAGAGATCTTGGCCGGAGTCTTCAGAGCTAACCTAACATCAGATGTTCCCA 1020
DB 1005 ACAGAGATCTTGGCCGGAGTCTTCAGAGCTAACCTAACATCAGATGTTCCCA 1064
QY 1021 CTGCTACGCTCGAAGCTTGTCTGGAACAGGCTGGGCAAGGTTCTGAAAGT 1080
DB 1065 CTGCTACGCTCGAAGCTTGTCTGGAACAGGCTGGGCAAGGTTCTGAAAGT 1124
QY 1081 GAAAGCTGGCTCCAGATGCTAAAGGCTTTCCAGAGACCGGCTCTCCGCCAGAA 1140
DB 1125 GAAAGCTGGCTCCAGATGCTAAAGGCTTTCCAGAGACCGGCTCTCCGCCAGAA 1184

QY 1141 GCACTGTTTCATGATTATTTGACGCCCTGCGCATCTCAGCTGATGCAAGCTTCTCTATGAG 1200
DB 1185 GCACTGTTTCATGATTATTTGACGCCCTGCGCATCTCAGCTGATGCAAGCTTCTCTATGAG 1244
QY 1201 GAGTCTTTGTTTACAGTTTCAGAGTGAAGCTTAAGCCAGAAATGATGACCTTTTGAGGC 1260
DB 1245 GAGTCTTTGTTTACAGTTTCAGAGTGAAGCTTAAGCCAGAAATGATGACCTTTTGAGGC 1304
QY 1261 TCCCTACCAAGAGTGCACCAAGCCAGCTTTAGCAAAATGCTGTGTA 1308
DB 1305 TCCCTACCAAGAGTGCACCAAGCCAGCTTTAGCAAAATGCTGTGTA 1352

RESULT 4
US-09-801-861-1
; Sequence 1, Application US/09801861
; Patent No. US20020119544A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01098
; CURRENT APPLICATION NUMBER: US/09/801,861
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ For Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2203
; TYPE: DNA
; ORGANISM: Human
US-09-801-861-1

Query Match 77.1%; Score 1009; DB 9; Length 2203;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1009; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTCAGAGCTGTGTGCAAAAGCTGTAACCTCTGATGCAAGCTGCTTAACCTTTGTTCA 60
DB 118 ATGGGTCAGAGCTGTGTGCAAAAGCTGTAACCTCTGATGCAAGCTGCTTAACCTTTGTTCA 177
QY 61 GAGGAGGCGAGGCAACAGCTGTCGAGAGAGTCAAGCTGAGACACGAGAGCTGAGTTC 120
DB 178 GAGGAGGCGAGGCAACAGCTGTCGAGAGAGTCAAGCTGAGACACGAGAGCTGAGTTC 237
QY 121 AAGCTAACAGACTTAAGAGATCATGTTCCATGACTTCAATTCACCCAGAGGACTT 180
DB 238 AAGCTAACAGACTTAAGAGATCATGTTCCATGACTTCAATTCACCCAGAGGACTT 297
QY 181 CAAGCTGCCCCGTCGCAAGATTCAGAGTAAAGGCCACGAGTAAAGATGTTGTTT 240
DB 298 CAAGCTGCCCCGTCGCAAGATTCAGAGTAAAGGCCACGAGTAAAGATGTTGTTT 357
QY 241 CAGGAAGAGATCTGAGCAGGGTTTCACTGAGAGAGAGCTCCCTTTTGGGAGAGCC 300
DB 358 CAGGAAGAGATCTGAGCAGGGTTTCACTGAGAGAGAGCTCCCTTTTGGGAGAGCC 417
QY 301 TCATCTTACTTGAATCTGGAAGAGCTGGGTGAAGGCTTTATGCAAGTTTCAAGGGG 360
DB 418 TCATCTTACTTGAATCTGGAAGAGCTGGGTGAAGGCTTTATGCAAGTTTCAAGGGG 477
QY 361 ATTAAGCAAAATTAATGACAACTATGAGCTTTAAAGTCACTGAGTGAATGAGAGAA 420
DB 478 ATTAAGCAAAATTAATGACAACTATGAGCTTTAAAGTCACTGAGTGAATGAGAGAA 537
QY 421 GAGTCCCATTTTACAGCTATCCGAAAGCTTCTCTGTAAGGGTTTGAACAATGCAAT 480
DB 538 GAGTCCCATTTTACAGCTATCCGAAAGCTTCTCTGTAAGGGTTTGAACAATGCAAT 597
QY 481 ATTGTCTCTCTGATGACATTAATCAACCAAGAGACATGACATTCGTTTGTGAATAC 540
DB 598 ATTGTCTCTCTGATGACATTAATCAACCAAGAGACATGACATTCGTTTGTGAATAC 657

QY 541 ATGCACACAGACTGGGCCAGATATATGTTTCAAGCATCCAGAGAGGCTTATCTCTATAT 600
DB 658 ATGCACACAGACTGGGCCAGATATATGTTTCAAGCATCCAGAGAGGCTTATCTCTATAT 717
QY 601 GTCCAGCTTTTATGTTTCACTTTTGGGGGCTGGGCTGATCATCCACCAACAACGTT 660
DB 718 GTCCAGCTTTTATGTTTCACTTTTGGGGGCTGGGCTGATCATCCACCAACAACGTT 777
QY 661 CTTCACAGGGACCTGAACCTTCAGAACTTATCATCATGCTACCTGGAGAGCTCAACTG 720
DB 778 CTTCACAGGGACCTGAACCTTCAGAACTTATCATCATGCTACCTGGAGAGCTCAACTG 837
QY 721 GCTGATTTTGTGTTGCTGGCCGGGCGCAAGTCCATTTCCAGCAGACATCTTCAGAGTC 780
DB 838 GCTGATTTTGTGTTGCTGGCCGGGCGCAAGTCCATTTCCAGCAGACATCTTCAGAGTC 897
QY 781 GTGACCTCTGGTACCGGCCCCCTGATGCTTGTGAGAGCAGTGAATATTCCTCTGAG 840
DB 898 GTGACCTCTGGTACCGGCCCCCTGATGCTTGTGAGAGCAGTGAATATTCCTCTGAG 957
QY 841 CTGACATATGAGGGGTGACAGCTGCAATCTTATGAAATGTTCCAGAGTCAACCTTTGTT 900
DB 958 CTGACATATGAGGGGTGACAGCTGCAATCTTATGAAATGTTCCAGAGTCAACCTTTGTT 1017
QY 901 CTTGGGCTTTCCAAATCTTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCT 960
DB 1018 CTTGGGCTTTCCAAATCTTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCT 1077
QY 961 ACAGAGATCTTGGCCGGAGTCTTCAAGCTAACCTAACATTCAG 1009
DB 1078 ACAGAGATCTTGGCCGGAGTCTTCAAGCTAACCTAACATTCAG 1126

RESULT 5
US-10-224-562-1
; Sequence 1, Application US/10224562
; Publication No. US20030022229A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01098DIV
; CURRENT APPLICATION NUMBER: US/10/224,562
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ For Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-224-562-1

Query Match 77.1%; Score 1009; DB 14; Length 2203;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1009; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTCAGAGCTGTGTGCAAAAGCTGTAACCTCTGATGCAAGCTGCTTAACCTTTGTTCA 60
DB 118 ATGGGTCAGAGCTGTGTGCAAAAGCTGTAACCTCTGATGCAAGCTGCTTAACCTTTGTTCA 177
QY 61 GAGGAGGCGAGGCAACAGCTGTCGAGAGAGTCAAGCTGAGACACGAGAGCTGAGTTC 120
DB 178 GAGGAGGCGAGGCAACAGCTGTCGAGAGAGTCAAGCTGAGACACGAGAGCTGAGTTC 237
QY 121 AAGCTAACAGACTTAAGAGATCATGTTCCATGACTTCAATTCACCCAGAGGACTT 180
DB 238 AAGCTAACAGACTTAAGAGATCATGTTCCATGACTTCAATTCACCCAGAGGACTT 297
QY 181 CAAGCTGCCCCGTCGCAAGATTCAGAGTAAAGGCCACGAGTAAAGATGTTGTTT 240
DB 298 CAAGCTGCCCCGTCGCAAGATTCAGAGTAAAGGCCACGAGTAAAGATGTTGTTT 357


```
Db 746 TCAGACTTTTCATGTTTCAACTTTTGCGGGGCGCTGCGTACATCCACCAACAGCTTC 805
Qy 662 TTCACAGGAGCCTGAAAACCTCAGAACTTACTCATGCTACCTGGAGAGCTCAAACTGG 721
Db 806 TTCACAGGAGCCTGAAAACCTCAGAACTTACTCATGCTACCTGGAGAGCTCAAACTGG 865
Qy 722 CTGATTTTGGTCTTTGCGCGGGGCGAAGTCCATTTCCAGCCGAGATACTCTTTCAGAAAGTCG 781
Db 866 CTGATTTTGGTCTTTGCGCGGGGCGAAGTCCATTTCCAGCCGAGATACTCTTTCAGAAAGTCG 925
Qy 782 TGACCTCTGATGATCCGGGCGCTGATGCTTTGGTGGAGGACCTGATATATTTCCCTGAGC 841
Db 926 TGACCTCTGATGATCCGGGCGCTGATGCTTTGGTGGAGGACCTGATATATTTCCCTGAGC 985
Qy 842 TGACATATGAGGCTGACGCTGATCTTTATGAAATGTTTCAAGGCTCAACCTTTGTTTC 901
Db 986 TGACATATGAGGCTGACGCTGATCTTTATGAAATGTTTCAAGGCTCAACCTTTGTTTC 1045
Qy 902 CTGGGGGTTTCCACATCCCTTGAACAGCTGAGAAATCTGGAGGCTGCTGGAGTCCCTA 961
Db 1046 CTGGGGGTTTCCACATCCCTTGAACAGCTGAGAAATCTGGAGGCTGCTGGAGTCCCTA 1105
Qy 962 CAGAGGATATCTTGGCGGAGTCTCCAGCTACTACTAATCCAGAAATGCT 1015
Db 1106 CAGAGGATATCTTGGCGGAGTCTCCAGCTACTACTAATCCAGAAATGCT 1159

RESULT 7
US-10-425-114-26244
; Sequence 26244, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26244
; LENGTH: 3210
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4119-079-cl_FLI
US-10-425-114-26244

Query Match 65.1%; Score 851.8; DB 16; Length 3210;
Best Local Similarity 99.7%; Pred. No. 1.4e-275;
Matches 864; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 443 GAGAGCTTCTCTCCGGAAGGTTGAACATGCCAATTTGTCTCCGTCATGACATTA 502
Db 59 GATTACTCTCTCTCCGGAAGGTTGAACATGCCAATTTGTCTCCGTCATGACATTA 118
Qy 503 TCCACACCAAGAGACACTGATCTGTTTGAATACAGACACAGACCTGGCCCACT 562
Db 119 TCCACACCAAGAGACACTGATCTGTTTGAATACAGACACAGACCTGGCCCACT 178
Qy 563 ATATGCTCAGATCCAGAGGCTTCACTCATATATGTCAGACTTTTCATGTTTCAAC 622
Db 179 ATATGCTCAGATCCAGAGGCTTCACTCATATATGTCAGACTTTTCATGTTTCAAC 238
Qy 623 TTTTGGGGGCGCTGGGTCATCCACCAACACAGTCTTTTCACAGGAGCTGAAACCTTC 682
Db 239 TTTTGGGGGCGCTGGGTCATCCACCAACACAGTCTTTTCACAGGAGCTGAAACCTTC 298
Qy 683 AGAAGTACTCATGATCAGTCACTGGAGAGCTCAAACTGCTGATTTTGGTCTTGGCGGAG 742
```

```
Db 299 AGAAGTACTCATGATCAGTCACTGGAGAGCTCAAACTGCTGATTTTGGTCTTGGCGGAG 358
Qy 743 CCAAGTCCATTCCTCCAGCCAGACATATCTTTAGAAAGTCGTGACCCCTGGTACCGGCCCC 802
Db 359 CCAAGTCCATTCCTCCAGCCAGACATATCTTTAGAAAGTCGTGACCCCTGGTACCGGCCCC 418
Qy 803 CTGATGCTTTGCTGGAGGCACTGATATTTCTCTGAGCTGAGACATATGAGGCTGACAGCT 862
Db 419 CTGATGCTTTGCTGGAGGCACTGATATTTCTCTGAGCTGAGACATATGAGGCTGACAGCT 478
Qy 863 GCATCTTTATGAAATGTTCCAGGCTTCAACCTTTGTTTCCTGGGTTTCCATTCCTTG 922
Db 479 GCATCTTTATGAAATGTTCCAGGCTTCAACCTTTGTTTCCTGGGTTTCCATTCCTTG 538
Qy 923 AACAGCTGAGAAATCTGGAGGCTGCTGGAGTCCCTAAGAGATATTTGGCGGAG 982
Db 539 AACAGCTGAGAAATCTGGAGGCTGCTGGAGTCCCTAAGAGATATTTGGCGGAG 598
Qy 983 TCTCCAGCTACTACTAATCAATCCAGAAATGCTTCCACTGCTTCAAGCTTTC 1042
Db 599 TCTCCAGCTACTACTAATCAATCCAGAAATGCTTCCACTGCTTCAAGCTTTC 658
Qy 1043 ATGTTGCTGGAACAGGCTGGGCGAGGTTCTGAGAGCTGAGACCTGGCTCCAGATGC 1102
Db 659 ATGTTGCTGGAACAGGCTGGGCGAGGTTCTGAGAGCTGAGACCTGGCTCCAGATGC 718
Qy 1103 TAAAGGCTTTTCCAGAGCCGCTCTCCGCCAGAGACATTTGTATGATTTTCA 1162
Db 719 TAAAGGCTTTTCCAGAGCCGCTCTCCGCCAGAGACATTTGTATGATTTTCA 778
Qy 1163 GCGCCCTGCTCATCTGAGCTGTACAGCTTCTGATGAGAGTCTTTGTTTACAGTTTCAG 1222
Db 779 GCGCCCTGCTCATCTGAGCTGTACAGCTTCTGATGAGAGTCTTTGTTTACAGTTTCAG 838
Qy 1223 GAGTAGGCTTAAAGCAG-AAATGCTGACCTTTGGCTCTACAGAAAGGTACACAC 1281
Db 839 GAGTAGGCTTAAAGCAGAAATGCTGACCTTTGGCTCTACAGAAAGGTACACAC 898
Qy 1282 CCAGCCAGTTTATGCAAAATGCTGTGA 1308
Db 899 CCAGCCAGTTTATGCAAAATGCTGTGA 925

RESULT 8
US-09-948-802-14
; Sequence 14, Application US/09948802
; Publication No. US20020172981A1
; GENERAL INFORMATION:
; APPLICANT: ROBINSON, KEITH E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; FILE REFERENCE: MNI-090
; CURRENT APPLICATION NUMBER: US/09/948,802
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/387,212
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-948-802-14

Query Match 36.8%; Score 481; DB 9; Length 481;
Best Local Similarity 100.0%; Pred. No. 6.1e-151;
Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 720 GCGTATTTTGGTCTTGGCCGCGGCAAGTCCATTTCCAGCAGACATATCTTCAAGAGT 779
Db 1 GCGTATTTTGGTCTTGGCCGCGGCAAGTCCATTTCCAGCAGACATATCTTCAAGAGT 60
```

QY 780 CGTGAACCTCTGTAACCGGCCCCCTGATGCTTGTGGAAGCACTGAATATTCCTCTGA 839
DB 61 CGTGAACCTCTGTAACCGGCCCCCTGATGCTTGTGGAAGCACTGAATATTCCTCTGA 120
QY 840 GCTGACATATGAGGAGTGCAGGCTGATCTTTATTTGAAATGTTCCAGGGTCAACCTTTGTT 899
DB 121 GCTGACATATGAGGAGTGCAGGCTGATCTTTATTTGAAATGTTCCAGGGTCAACCTTTGTT 180
QY 900 TCTGTGGGTTTCCAACTCTTGAACAGCTGGAAGAAATCTGGAGAGTCTGGAGATCCC 959
DB 181 TCTGTGGGTTTCCAACTCTTGAACAGCTGGAAGAAATCTGGAGAGTCTGGAGATCCC 240
QY 960 TACAGAGATTAATGAGGAGTCTCCAAAGTACTTAATCAATCCAGATGTTCTCC 1019
DB 241 TACAGAGATTAATGAGGAGTCTCCAAAGTACTTAATCAATCCAGATGTTCTCC 300
QY 1020 ACTGCTACGCTCGAAGCTTCAATGTTGTCTGAAACAGGCTGGGAGGTTCTGTAAC 1079
DB 301 ACTGCTACGCTCGAAGCTTCAATGTTGTCTGAAACAGGCTGGGAGGTTCTGTAAC 360
QY 1080 TGAAGACCTGAGCTCCAGAGTCTAAAGGCTTTCCAGAGACCGGCTCTCCGCCAGGA 1139
DB 361 TGAAGACCTGAGCTCCAGAGTCTAAAGGCTTTCCAGAGACCGGCTCTCCGCCAGGA 420
QY 1140 AGCACTTGTATGATTAATTTAGAGGCTCTGCAATCTCAGCTGTACAGCTTCTGATGA 1199
DB 421 AGCACTTGTATGATTAATTTAGAGGCTCTGCAATCTCAGCTGTACAGCTTCTGATGA 480
QY 1200 G 1200
DB 481 G 481

RESULT 9
US-10-121-925-14
Sequence 14, Application US/10121925
Publication No. US20030104505A1
GENERAL INFORMATION:
APPLICANT: ROBINSON, KEITH E.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
FILE REFERENCE: MNT-090
CURRENT APPLICATION NUMBER: US/10/121,925
PRIOR FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: US/09/948,802
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 09/387,212
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 14
LENGTH: 481
TYPE: DNA
ORGANISM: Homo sapiens
US-10-121-925-14

Query Match 36.8%; Score 481; DB 15; Length 481;
Best Local Similarity 100.0%; Pred. No. 6,1e-151;
Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 720 GGCTGATTTTGTCTTGGCCCGGACCAATGTCATTTCCAGCCAGACATCTTTCAGAGT 779
DB 1 GGCTGATTTTGTCTTGGCCCGGACCAATGTCATTTCCAGCCAGACATCTTTCAGAGT 60
QY 780 CGTGAACCTCTGTAACCGGCCCCCTGATGCTTGTGGAAGCACTGAATATTCCTCTGA 839
DB 61 CGTGAACCTCTGTAACCGGCCCCCTGATGCTTGTGGAAGCACTGAATATTCCTCTGA 120
QY 840 GCTGACATATGAGGAGTGCAGGCTGATCTTTATTTGAAATGTTCCAGGGTCAACCTTTGTT 899
DB 121 GCTGACATATGAGGAGTGCAGGCTGATCTTTATTTGAAATGTTCCAGGGTCAACCTTTGTT 180
QY 900 TCTGTGGGTTTCCAACTCTTGAACAGCTGGAAGAAATCTGGAGAGTCTGGAGATCCC 959

DB 181 TCTGTGGGTTTCCAACTCTTGAACAGCTGGAAGAAATCTGGAGAGTCTGGAGATCCC 240
QY 960 TACAGAGATTAATGAGGAGTCTCCAAAGTACTTAATCAATCCAGATGTTCTCC 1019
DB 241 TACAGAGATTAATGAGGAGTCTCCAAAGTACTTAATCAATCCAGATGTTCTCC 300
QY 1020 ACTGCTACGCTCGAAGCTTCAATGTTGTCTGAAACAGGCTGGGAGGTTCTGTAAC 1079
DB 301 ACTGCTACGCTCGAAGCTTCAATGTTGTCTGAAACAGGCTGGGAGGTTCTGTAAC 360
QY 1080 TGAAGACCTGAGCTCCAGAGTCTAAAGGCTTTCCAGAGACCGGCTCTCCGCCAGGA 1139
DB 361 TGAAGACCTGAGCTCCAGAGTCTAAAGGCTTTCCAGAGACCGGCTCTCCGCCAGGA 420
QY 1140 AGCACTTGTATGATTAATTTAGAGGCTCTGCAATCTCAGCTGTACAGCTTCTGATGA 1199
DB 421 AGCACTTGTATGATTAATTTAGAGGCTCTGCAATCTCAGCTGTACAGCTTCTGATGA 480
QY 1200 G 1200
DB 481 G 481

RESULT 10
US-10-153-242-3
Sequence 3, Application US/10153242
Publication No. US20030166217A1
GENERAL INFORMATION:
APPLICANT: Reinhard, Christoph
APPLICANT: Pot, David
APPLICANT: Kassam, Alcaef
APPLICANT: Marensbach, Taasha
APPLICANT: Williams, Lewis T.
TITLE OF INVENTION: HUMAN CYCLIN-DEPENDENT KINASE (hPFTAIRE)
FILE REFERENCE: PP-01429,103/200130,445D1
CURRENT APPLICATION NUMBER: US/10/153,242
CURRENT FILING DATE: 2002-05-22
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 2467
TYPE: DNA
ORGANISM: mouse
US-10-153-242-3

Query Match 33.5%; Score 438; DB 15; Length 2467;
Best Local Similarity 65.8%; Pred. No. 4,9e-136;
Matches 636; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

QY 277 AAGACCTCCCTTTTGGGAGCACTCATCTTATTTGAATTTGGAAGCTGGGTGAAGC 336
DB 353 ACGAGTCCCAATTTTGAAGAGTCACTCATTCGAAAACTGGAAGAACTGGGGAGGA 412
QY 337 TCTTATGCAACGTTTACAGGGAGTTAGCAGATTAATGACAACTAGTGCCTTTAAAG 396
DB 413 TCTTATGCAACGTTTACAGGGAGTTAGCAGATTAATGAGAACTAGTGCCTTTAAAG 472
QY 397 GTCATCAGCATGAAGGAGAGAGAGTCCATTATTAAGTATCGAAGACTTCTCTC 456
DB 473 GTCATCAGCATGAAGGAGAGAGAGTCCATTATTAAGTATCGAAGACTTCTCTC 532
QY 457 CTGAAGGTTTGAACATGCCAATATTTGCTCTGCAATGATATCCACCAAGAG 516
DB 533 TTGAAGAGCTTAAGACAGGCCAATCGTGTGCTTCAAGACATCATCACTAAGGA 592
QY 517 ACATGCACTTGTGTTTGAATATCATGACACAGACCTGGCCATATATGCTCAGCT 576
DB 593 ACCCTGACCTTGTGTTTGAATATCATGACACAGACCTGGCCATATATGCTCAGCT 652
QY 577 CGAAGAGGCTTCAATCCATATATGTCAGACTTTCATGTTTCACTTGGGGGCTG 636
DB 653 CTTGAAGACTCATTCAGATTAATGTAAGTGTGTTTATTTTCACTGCTGGAGACTG 712

QY 637 GCCTACATCCACCAACAGCTTCTTCAAGAGACCTGAACCTCAAGAACTTACTGATC 636
DB 713 TCTTACATCCACAGCGCTTATTTTGGACAGAGACCTGAACCGCAGAACTTCTCATC 772
QY 697 AGTCACTGGAGAGCTCAACCTGGCTGATTTGGTCTTTCGCCGCCAAGCTCATTTCC 756
DB 773 AGGATTCGGGGAGTTGAAGCTGGACGATTTGGCTGGACAGAAATCCGCTCT 832
QY 757 AGCCAGACATATCTTTCAGAAAGTCGACCCCTCTGATCCGCCCTGATGCTTTCCTG 816
DB 833 AGCCAGACATATCTTTCAGAAAGTTGATTCCTTGGATGACAGCTCCAGATGTTCTTCG 892
QY 817 GGAGCCACTGAATTTCTCTGAGCTGACATATGGGCTGACAGGCTGATCTTTATTGAA 876
DB 893 GGGCTTACAGAAATTTCCACTGCTGACATGTGGGAGTTGGCTGTATCTTCGTTGAG 952
QY 877 ATGTTCCAGGGTCAACCTTTGTTCTCGGGGTTTCCAACTCTTGAACAGCTGAGAAA 936
DB 953 ATGATCCAGAGAGTTGCTGGCTTTCCAGGATGAAGAACTTACAGATCACTTGAACGG 1012
QY 937 ATCTGGAGGTGCTGGAGATCCCTACAGAGATATTTGGCCGGAGTCTTCAAGCTACT 996
DB 1013 ATATTTCTGGTTCTTGGAAACCGAAATGAGACACGTTGGCTGATTTCTTTAACCA 1072
QY 997 AACTCAATCCAGAAATGTTCCCACTGCTTACCGCTTCAAGCTTCACTGTTCTGGAAC 1056
DB 1073 CATTTTAAGCAAAACGCTTTTACCGTGTACAACTTAAAGCCCTTGAACAGATGGAAT 1132
QY 1057 AGGCTGGGAGGTTCTCTGAAGCTGAAGACCTGAGCTCCAGATGCTTAAAGGCTTTCC 1116
DB 1133 AACTCAAGTATGTAAATATGCTGAAGACTTGGCTTCCAGCTTCTCCAGTGTCCCA 1192
QY 1117 AGAAGCCGCTCTCCGCCAGAAAGCACTTGTTCATGATTAATTTCAAGCCCTGCCATCT 1176
DB 1193 AAGAAAGGCTATCAGACAGAGCCGCTTGAACATGATTAATTTCAAGCATCTGCCCTCA 1252
QY 1177 CAGCTGATCAGCTTCTCTGATGAGAGCTCTTTTTCACGTTTTCAGAGCTGAGAGCTAAAG 1236
DB 1253 CGGCTATGGAGCTGACTGATGATGCTTCTTATTTTACCGTCCAAATGATGAGATTGCAA 1312
QY 1237 CCAGAA 1242
DB 1313 CCAGAA 1318

RESULT 11
US-10-153-242-1
; Sequence 1, Application US/10153242
; Publication No. US20030166217A1
; GENERAL INFORMATION:
; APPLICANT: Reinhard, Christoph
; APPLICANT: Pot, David
; APPLICANT: Kassam, Altaf
; APPLICANT: Marenbach, Tasha
; APPLICANT: Williams, Lewis T.
; TITLE OF INVENTION: HUMAN CYCLIN-DEPENDENT KINASE (hPPTAIRE)
; FILE REFERENCE: PP-01429.103/200130.445D1
; CURRENT FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1272
; TYPE: DNA
; ORGANISM: human
US-10-153-242-1

Query Match 33.5%; Score 437.8; DB 15; Length 1272;
Best Local Similarity 65.1%; Pred. No. 3.9e-136;
Matches 646; Conservative 0; Mismatches 347; Indels 0; Gaps 0;
QY 280 AGCTTCCTTTTGGGAGGCTCATCTTACTTGAACCTGAGAAAGCTGGGTGAAGGCTCT 339

DB 238 AGTCCCAATTTTGAAGAAAGCTGATCTATTTGAAGAAAGCTGAGAAAAGCTAGGGAGGATCT 297
QY 340 TATGCAAGATTTTACAAAGGGATTAGCAGATTAATATGACAACTAGTGGCTTTAAAGTC 399
DB 298 TATGCTACGATATACAAAGGGAAGAAAGGATTAATGGAAGTTGTAGCTCTGAAGGTG 357
QY 400 ATCAGATGAATGACAGAGAGAGAGTCCCATTTTACAGCTATCCAGAAAGCTTCTCTG 459
DB 358 ATCAGGCTCAGAGAAAGAAAGAGGACACTTTTACAGCTATCAGGGAAAGCTTCTCTTTA 417
QY 460 AAGGTTTGAACATGCCCATAATGCTCCGATGATGATATATCCACCAAGAGAGA 519
DB 418 AAGGACTAAGCATGCTAATAGTCTTCACTTCAATGATATATATCAAGAGAGAG 477
QY 520 CTGACATGTTGTTTGAATATACATGACACAGACTGAGCCAGATATATGCTCAGATCCA 579
DB 478 CTGACACTTGTGTTGAATATGATGACACTGATTTATGCTACATGACATGACAAAGCACT 537
QY 580 GGAAGGCTTCACTCTCATATATGTCAGACTTTTCATGTTTCAACTTTGCGGGGCTGGCG 639
DB 538 GGGGGGCTGATCCAGATATATGTAAGTTGTTTATTTTCAAGTGTCTGAGAGTCTGTCT 597
QY 640 TATATCCACCAACAAACAGTCTTCTTCAAGGGAAGCTGAAAGCTCAAGATCTTACTCATGCT 699
DB 598 TATATCCACCAAGGTTATATTTTGAACAGAGACTGAAACCAAGAACTTCTGATCACT 657
QY 700 CACTGGGAGAGCTGAACTGAGCTATTTTGTCTTGGCTTGGCCGGGCAAGTCCAATCCAGAC 759
DB 658 GACAGGGGGAGTTAAAGCTGGCAATTTTGGCTTTGCAAGAGCAAAATCCGCTCCAGAC 717
QY 760 CAGACATATCTTTCAGAAAGTGTGACCTTGTGATCCGGCCCTGATGCTTGTGGGA 819
DB 718 CACATATCTTCAAGAAAGTGTGATCTTGTGTGACAGCTTCCAGATGCTCTTCAAGGC 777
QY 820 GGCACGAAATATCTCTGAGCTGACATATGAGGAGTGGAGCTGATCTTTATTTGAATG 879
DB 778 TCAAGAAATATTTCAACCTGCTTGAACATGTGGGAGTATGATTTGATTTGAAATG 837
QY 880 TTCCAGGCTCAACTTTGTTTCTGGGGTTTCCAACTATCTTTGAACAGCTGAGAGAAATC 939
DB 838 ATCCAAAGGAGTGTGCTGTTTCCAGGAATGAAAGCAATTCAGAGTCACTTGAAGGAAT 897
QY 940 TGGAGGTGCTGGAGTCCCTTCAAGAGATATTTGGCCGGAGTCTCAAGCTTACTTAC 999
DB 898 TTTCTGTTCTTGAACCAACAAATGAGACAATGGCCCTGAGTTCATCTTTTACCAAT 957
QY 1000 TACAATCCAGAAATGTTCCCACTGCTTACGCTCGAAGGCTTCAATGTTGTCTGAAACAG 1059
DB 958 TTTAAAGCCAGAAAGCTTTTACCTGTATCAGCTTTAAAGCTTAAAGCATGAAATGAG 1017
QY 1060 CTGGGCAAGGTTCTGAAGCTGAAGACCTGCTCCCAAGTCTTAAAGGCTTTTCCAGA 1119
DB 1018 CTCAGCTATGTGAACCATGACAGAGACTGGCTCCAAAGCTCTTCAATGTTCCCAAG 1077
QY 1120 GACCGGCTCTCGCCCAAGAGCACTTGTATATATTTTCAAGGCTTCTGCACTCAG 1179
DB 1078 AACCAACTGTGGCAACAGCTGCTTGAAGCAAGGATTTTATGACACTGCGCAAG 1137
QY 1180 CTGTACCAAGCTTCTGATGAGAGAGTGTGTTTACAGTTTCAAGAGTGAAGGCTAAAGCA 1239
DB 1138 CTATGGAGACTACCAAGCATGCTTCTTATTTTACTGTGCCAAATGTATGATGACACCA 1197
QY 1240 GAATGTGTGACCTTTTGGCTCTTACAGAAA 1272
DB 1198 GAAGCTGAGAAAGCATGCGGGCTTTTGGGAAA 1230

RESULT 12
US-10-153-242-13
; Sequence 13, Application US/10153242
; Publication No. US20030166217A1
; GENERAL INFORMATION:

```

? APPLICANT: Reinhard, Christoph
? APPLICANT: Pot, David
? APPLICANT: Kaasam, Altaf
? APPLICANT: Marendach, Tasha
? APPLICANT: Williams, Lewis T.
? TITLE OF INVENTION: HUMAN CYCLIN-DEPENDENT KINASE (HPPTAIRE)
? FILE REFERENCE: PP-01429.103/200130.445D1
? CURRENT APPLICATION NUMBER: US/10/153,242
? CURRENT FILING DATE: 2002-05-22
? NUMBER OF SEQ ID NOS: 13
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 13
? LENGTH: 1549
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-10-153-242-13

Query Match      33.3%   Score 436.2, DB 15, Length 1549,
Best Local Similarity 65.0%   Pred. No. 1,5e-135,
Matches 645, Conservative 0, Mismatches 346, Indels 0, Gaps 0

```

Db 886 GACACGGGGAGTTAAAGCTGGACGATTTGCGTCTTCCAGAGCAAAATCGTCCTAGC 945
Qy 760 CACACATCTCTTCAAGAGTGTGAACCTCTGGTACCGGCCCTGATGCTTTGCTGGGA 819
Db 946 CACACATCTCTTCAAGAGTGTGAACCTCTGGTACCGGCCCTGATGCTTTGCTGGGA 1005
Qy 820 GCCATCTGATTTCTCTGAGCTGACATATGGGGTGCAGAGCTGCATCTTTATTTGAATG 879
Db 1006 TCACAGAAATATTCACCTGCTGTGACATGTGGGAGTAGGTGCACTTTGTTGAATG 1065
Qy 880 TTCAGGGTCAACCTTTGTTCTCTGGGGTTTCCAACTCTTGAACAGCTGAGAAATC 939
Db 1066 ATCCAGAGATCTGCTGTTTCCAGAAATGAAGACATTTAGATCACTTGAACGAAT 1125
Qy 940 TGGAGGTGTGGAGTCCCTTACAGAGATATCTGGCCGGAGTCTTCAAGCTACCTAAC 999
Db 1126 TTTCTGGTCTTGGAAACCAATAGAGACACATGGCTGGAGTTGATCTTACCAT 1185
Qy 1000 TACAATCCAGAAATGTTCCTCCATGCTACGCTCGAAGCTTCTATGTTCTTGGAAACAG 1059
Db 1186 TTTAAGCCAGAAACGCTTTTACCTGTACAGCTTAAACCTTAGACAAAGCATGAAATAG 1245
Qy 1060 CTGGGAGGGTCTCTGAAGCTGAAGACTGGCTCCAGATGCTAAAGGCTTTCCAGA 1119
Db 1246 CTGAGTATGTGAACATGAGAGAGAGCTGGCTCCAGAGCTCTTACATGTTCCCAAG 1305
Qy 1120 GACCGGTCTCCGCCCAAGAGACATTTTCATGATTTATTTAGCCGCTGCATCTCAG 1179
Db 1306 AACAGACTGTGGCAGAGGCTGCTTGAAGCCAGAGATTTTATGATGACCTGCGCACAG 1365
Qy 1180 CTGTACAGCTCTGATGAGAGTCTTTGTTACAGTTTCAGAGATGAGAGCTTAAGCCA 1239
Db 1366 CTATGGGAATCAACGACATGCTTCTATTTTACGTCTCCAAATGAGATTGCAACCA 1425
Qy 1240 GAAATGTGACCTTTTGGCTCTTACAGAAA 1272
Db 1426 GAAGCTGAGAAAGCATGGGGCTTTGGGAAA 1458

RESULT 14
US-10-723-860-5096
; Sequence 5096, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5096
; LENGTH: 5161
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4985)..(5005)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-5096

Query Match 33.3%; Score 436.2; DB 18; Length 5161;
Best Local Similarity 65.0%; Pred. No. 36-135;
Matches 645; Conservative 0; Mismatches 348; Indels 0; Gaps 0;
Qy 280 AGCTTCCTTTGGGAGCTTCATCTTATCTTGAACCTTGAGAGAGCTGGGTGAGAGCTCT 339

Db 466 AGTCCCAATTTTGAAAAGCTGACTCATATGAAAAGCTGGAAAAGCTAGGGAGAGATCT 525
Qy 340 TATGCAAGATTTTACAAGGGGATTTAGCAGAAATTAATGACAACTAGTGGCTTTAAAGTC 399
Db 526 TATGCTACAGTATTAACAAGGAGAAAAGCAGATTAATGGAAGTTGTAGCTCTGAAGGTG 585
Qy 400 ATCAGATGAATGACAGAGAGAGATCCATTTTACAGATATCCGGAAGCTTCTCCCG 459
Db 586 ATCAGGCTCAGAGAAAGAAAGAGACACTTTCACAGCTATCAGGAGAGCTTCTTTTA 645
Qy 460 AAGGTTTGAACATGCCAATATGTGCTCTGCATGACATATCCACCAAGAGACA 519
Db 646 AAGAGCTAAACATGCTAATGATGCTACTTCAATGATCATCATCATCCAGAGAGAG 705
Qy 520 CTGACATTCGTTTTGATATACATGACACAGACCTGGCCAGATATGTCTCAGATCCA 579
Db 706 CTGACACTGTGTTGAAATATGTGACACTGATTTATGCTAGATAGACAGACACCT 765
Qy 580 GAGGGCTTCAATCCATATATGTCAGACTTTTCAATGTTTCACTTTGGGGGGCTGGGG 639
Db 766 GAGGGGCTGCATCCAGATATATGTAAGTTGTTTATTTTCAATGCTGCGAGGTCTGT 825
Qy 640 TACATCCACCAACACAGCTTCTTACAGAGAGCTGAAACCTCAGAACTTACTCATAGT 699
Db 826 TACATCCACCAACAGTTATATTTTGCACAGAGAGCTGAAACCAAGAACCTTCTGATCAGT 885
Qy 700 CACCTGGAGAGCTCAAACTGCTGATTTTGGTCTTGGCCGGGCCAAGTCCATTCCAGC 759
Db 886 GACAGGGGGAGTTAAAGCTGACAGATTTGCTTGGCAGAGCAAAATCCCTCTCTAGC 945
Qy 760 CACACATCTCTTCAAGAAATGTGACCTCTGCTGACCGGCCCTGATGCTTTGCTGGGA 819
Db 946 CACACATCTCTCAAGAAATGTGTTACCTTGTGTACAGACTCCAGATGCTCTTACAGC 1005
Qy 820 GCCATGATATTTCTCTGAGCTGACATATGGGGTGCAGGCTGACATCTTTATGAATG 879
Db 1006 TCACAGAAATTTTCAACCTGCTGCTTGAATGTGGGAGTATGATCTTTGTTGAATG 1065
Qy 880 TTCAGAGGTCAACCTTTGTTTCTTGGGGTTTCCAACTTCCTTGAACAGCTGAGAAATC 939
Db 1066 ATCCAGAGATTTGCTGCTTTTCCAGAAATGAAGACATTCAGAGATCACTTGAACGAATA 1125
Qy 940 TGGGAGGTGCTGGAGTCCCTTACAGAGATTAATGGCCGGAGTCTCCAGCTTACTTAAC 999
Db 1126 TTTCTGTTCTTGAACACCAAAATGAGACATATGCTTGAAGTCTTCTTACCACT 1185
Qy 1000 TACAATCCAGAAATGTTCCTCCATGCTTACAGCTTCAAGCTTCATGTTGTGAAACAG 1059
Db 1186 TTTAAGCCAGAAACGCTTTTACCTGTACAGCTTAATAAACCTTAGACAAAGCATGAATAAG 1245
Qy 1060 CTGGGCAAGGTTCCTGAAGCTGAAGACCTGGCTCCAGATGCTTAAGGCTTTCCAGA 1119
Db 1246 CTGACCTATGTGAACCATGACAGAGAGCTGGCTTCAAGCTCTTCAATGTTCCCAAG 1305
Qy 1120 GACCGGTCTCCGCCCAAGAGACATTTTCATGATTTATTTAGAGCCCTGCTCATCTCAG 1179
Db 1306 AACAGACTGTGGCAGAGAGCTGCTTGAAGCAGAGATTTTATGAGACTGCGCCAGAG 1365
Qy 1180 CTGTACCAAGCTTCTGATGAGAGTCTTTGTTTACAGTTTTCAGAGTGAAGCTTAAGCCA 1239
Db 1366 CTATGGGAATCAACGACATGCTTCTATTTTACGTCTCCAAATGAGATTGCAACCA 1425
Qy 1240 GAAATGTGACCTTTTGGCTCTTACAGAAA 1272
Db 1426 GAAGCTGAGAAAGCATGGGGCTTTGGGAAA 1458

RESULT 15
US-10-296-115-483
; Sequence 483, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc

TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
 FILE REFERENCE: 784dCT
 CURRENT APPLICATION NUMBER: US/10/296,115
 CURRENT FILING DATE: 2002-11-18
 PRIOR APPLICATION NUMBER: US09/488,725
 PRIOR FILING DATE: 2000-01-21
 PRIOR APPLICATION NUMBER: US09/552,317
 PRIOR FILING DATE: 2000-04-25
 NUMBER OF SEQ ID NOS: 1478
 SEQ ID NO 483
 LENGTH: 3024
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-296-115-483

Query Match 22.9%; Score 299.8; DB 16; Length 3024;
 Best Local Similarity 58.7%; Pred. No. 2.3e-89;
 Matches 577; Conservative 0; Mismatches 397; Indels 9; Gaps 3;

QY 289 TTTGGGCGAGCTCATCTTACTTGAACCTTGAGAGCTGGGTGAAGGCTTTATGCGACA 348
 DB 301 TTTGGGAACTGGAACATACGTAACCTGGAACCTGGAGAGGCACTATGCGACA 360
 QY 349 GTTTCAGAGGGATTAGAGAAATGGAACAAGTGGCTTTAAAGTCATCAGCATG 408
 DB 361 GTCTTCAAGGGGCGAGCAAACTGACGGAACCTTGTGACCTGAAAGATCCGGCTG 420
 QY 409 AATGCAAGAGAGAGTCCCATTTACAGTATCCGAGAGCTTCTCTGAAAGGTTTG 468
 DB 421 GAGCAGAGAGAGAGGCGCGCTGCACTGCAATCCGAGAGGTGTCTGTGAAGAACCTG 480
 QY 469 AAACATGCCAATATTGTGCTCTGTCATGACATATCCACACCAAGAGACACTGACATTG 528
 DB 481 AAGCAGCGCAATATTGTGACCTGTCATGACCTATCCACAGATGATGCTCCATCACCCTG 540
 QY 529 GTTTTGAATATGACACACAGACCGGCGCCAGTATATGATCAGATCCAGAGGCGTT 588
 DB 541 GTGTTTGAATGACCTGACAGTACGTAAGCAATATGATGACCACTGTGGAACTTCATG 600
 QY 589 CATCTCATATATGTCAGACTTTTCATGTTTCAACTTTTGGGGGCTGGCGTACATCAC 648
 DB 601 AGCATGCAACAAGTCAAGATTTTTCATGTTTCAAGTGTCTCGGGGCTCGGCTATGTCAC 660
 QY 649 CACCAACAGCTTCTTCAACAGGACCTGAAACCTCAGAACTTACTATCATGCTCCTGGGA 708
 DB 661 CACCGCAAGATCTGCAACGGGACCTGAAGCCCAAGAACCTGCTCATCAACGAGAGGGG 720
 QY 709 GAGCTCAACCTGCTGATTTTGTGCTTGGCCGGGCGCAATGCCATGCCAGCAGACATAC 768
 DB 721 GAGCTGAAGCTGCGCACTTTTGGCTGGCCAGGCGCAAGTCAAGTGCCTCAAAAGCTTAC 780
 QY 769 TCTTCAGAGTGTGACCTCTGTGTACCGGCGCCCTGATGCTTGTGGAGGCACTGAA 828
 DB 781 TCCATGAGGTGTGACCTGTGTGTACAGGCGCCCGGATGTGTGCTGGGATCCAGAG 840
 QY 829 TATTCCTGAGCTGACATATGAGGCTGACAGGCTGATCTTTAATGAAATGTTCCAGGGT 888
 DB 841 TACTCCACCCCATGATATGTGGGGCGTGGGCTGATCACAAGATGGGCAAGGG 900
 QY 889 CAACCTTGTGTTTCTGGGGTTTCCACATCTTTGAAACAGCTGGAAGAAATCTGGAGGTG 948
 DB 901 AGGCGCCCTTCTCC--GGGCTCCACAGTCAAGAGAGAGCTGCAAAATCAATGCGCTC 957
 QY 949 CTGGAGTCCCTTACAGAGGATCTTGGCGGGAGTCTCCAGGCTACTAATCAATCA 1008
 DB 958 CTCGGGACCCCAAGAGAGAGAGTGTGCGGGGTGACCGCTTCTGTAGTTCCGCAAC 1017
 QY 1009 GAATGTTCCACTGCTTACGCTTCAAGGCTTCAATGTTGTGAAACAGGCTGGGCAAG 1068
 DB 1018 TACAGCTTCC--CTGTCTACTCTCCGCAAGC---GCTCATCAACCAAGCGCCAGGTTG 1071
 QY 1069 GTTCTGAAGCTGAAGACTGCGCTCCAGATGCTAAAGGCTTTCCAGAGACCGCGTC 1128

DB 1072 GATACGAGATGCGATCCACCTCTGAGCAGCTGTCTGTATGAATCCAAAGATCGCATG 1131
 QY 1129 TCCGCCAGAGAGCACTTGTATATTAATTTCAGCGCCCTGCACTTCAAGTGTACCG 1188
 DB 1132 TCAGCAGAGGCTGCGCTGAGTCACTTCCGCTCTGAGAGAGGTGTGACCGAG 1191
 QY 1189 CTTCGATGAGAGTCTTGTATAGTTTCAAGTGTGAGGCTAAAGCCAGAAATGTGT 1248
 DB 1192 CTGGAAGACACTGCTCCATCTTCTCCTGAAGAGATCCAGCTCCAGAGAGCCAGGC 1251
 QY 1249 GACCTTTGGCCTCTTACAGAA 1271
 DB 1252 TACGAGGCTTGGCTTCCAGCA 1274

Search completed: December 27, 2004, 12:15:31
 Job time : 424.082 sec

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2004, 17:41:01, Search time 67.1305 Seconds
(without alignments)
13849.312 Million cell updates/sec

Title: US-10-786-065-4

Perfect score: 1308

Sequence: 1 atgggtcgaagctgtgtgc.....agttacgaatgcctgga 1308

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/prodata/1/ina/5C_COMB.seq:*

4: /cgn2_6/prodata/1/ina/5D_COMB.seq:*

5: /cgn2_6/prodata/1/ina/5E_COMB.seq:*

6: /cgn2_6/prodata/1/ina/5F_COMB.seq:*

7: /cgn2_6/prodata/1/ina/5G_COMB.seq:*

8: /cgn2_6/prodata/1/ina/5H_COMB.seq:*

9: /cgn2_6/prodata/1/ina/5I_COMB.seq:*

10: /cgn2_6/prodata/1/ina/5J_COMB.seq:*

11: /cgn2_6/prodata/1/ina/5K_COMB.seq:*

12: /cgn2_6/prodata/1/ina/5L_COMB.seq:*

13: /cgn2_6/prodata/1/ina/5M_COMB.seq:*

14: /cgn2_6/prodata/1/ina/5N_COMB.seq:*

15: /cgn2_6/prodata/1/ina/5O_COMB.seq:*

16: /cgn2_6/prodata/1/ina/5P_COMB.seq:*

17: /cgn2_6/prodata/1/ina/5Q_COMB.seq:*

18: /cgn2_6/prodata/1/ina/5R_COMB.seq:*

19: /cgn2_6/prodata/1/ina/5S_COMB.seq:*

20: /cgn2_6/prodata/1/ina/5T_COMB.seq:*

21: /cgn2_6/prodata/1/ina/5U_COMB.seq:*

22: /cgn2_6/prodata/1/ina/5V_COMB.seq:*

23: /cgn2_6/prodata/1/ina/5W_COMB.seq:*

24: /cgn2_6/prodata/1/ina/5X_COMB.seq:*

25: /cgn2_6/prodata/1/ina/5Y_COMB.seq:*

26: /cgn2_6/prodata/1/ina/5Z_COMB.seq:*

27: /cgn2_6/prodata/1/ina/5A_COMB.seq:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	1308	100.0	1308	US-09-801-861-4	Sequence 4, Appl1
2	1308	100.0	1308	US-10-224-562-4	Sequence 4, Appl1
3	1009	77.1	2203	US-09-801-861-1	Sequence 1, Appl1
4	1009	77.1	2203	US-10-224-562-1	Sequence 1, Appl1
5	481	36.8	481	US-09-387-212-14	Sequence 14, Appl1
6	481	36.8	481	US-09-948-802-14	Sequence 14, Appl1
7	437	33.5	2467	US-09-206-344A-3	Sequence 3, Appl1
8	437	33.5	2467	US-09-206-344A-1	Sequence 3, Appl1
9	436	33.3	1549	US-09-206-344A-13	Sequence 13, Appl1
10	296	22.6	2868	US-09-620-312D-53	Sequence 53, Appl1
11	291	22.3	1242	US-09-016-434-1283	Sequence 1283, Ap
12	216	16.5	1089	US-08-464-517-37	Sequence 37, Appl1
13	216	16.5	1089	US-08-464-517-37	Sequence 37, Appl1
14	216	16.5	1089	US-08-463-772-37	Sequence 37, Appl1
15	216	16.5	1089	US-08-463-772-37	Sequence 37, Appl1
16	216	16.5	1089	US-08-463-772-37	Sequence 37, Appl1
17	211	16.2	999	US-09-248-796A-4296	Sequence 4296, Ap
18	203	15.5	1161	US-09-023-655-1373	Sequence 1373, Ap
19	203	15.5	1161	US-09-023-655-1373	Sequence 1373, Ap
20	191	14.6	1476	US-08-869-106-1	Sequence 1, Appl1
21	191	14.6	1476	US-09-338-125-1	Sequence 1, Appl1
22	191	14.6	1476	US-09-266-225D-13	Sequence 13, Appl1
23	191	14.6	1635	US-09-417-197-112	Sequence 112, App
24	191	14.6	1635	US-09-417-197-112	Sequence 112, App
25	154	11.8	53332	US-09-801-861-3	Sequence 3, Appl1
26	154	11.8	53332	US-10-224-562-3	Sequence 3, Appl1
27	154	11.8	53332	US-10-224-562-3	Sequence 3, Appl1

ALIGNMENTS

28	145.4	11.1	278	4	US-09-016-434-333	Sequence 333, App
29	145.2	11.1	1050	4	US-09-220-132-3	Sequence 3, Appl1
30	145.2	11.1	1825	4	US-09-919-035-115	Sequence 115, App
31	136.2	10.4	903	2	US-08-874-347-9	Sequence 9, Appl1
32	136.2	10.4	903	4	US-09-093-522-9	Sequence 9, Appl1
33	126.4	9.7	993	4	US-09-248-796A-4323	Sequence 4323, Ap
34	126.4	9.7	1002	1	US-08-463-090B-3	Sequence 3, Appl1
35	121.4	9.3	1424	3	US-08-403-634-3	Sequence 3, Appl1
36	121.4	9.3	1424	3	US-08-913-441B-3	Sequence 3, Appl1
37	114.6	8.8	2471	4	US-09-915-060A-3	Sequence 3, Appl1
38	114.6	8.8	3883	4	US-09-620-312D-792	Sequence 792, Appl1
39	113.4	8.7	3505	1	US-07-660-465-11	Sequence 1, Appl1
40	112	8.6	945	4	US-09-671-050-11	Sequence 11, Appl1
41	112	8.6	972	4	US-09-671-050-5	Sequence 5, Appl1
42	112	8.6	1041	4	US-09-671-050-9	Sequence 9, Appl1
43	110	8.4	1701	4	US-09-671-050-3	Sequence 3, Appl1
44	110	8.4	1701	4	US-09-411-628-5	Sequence 5, Appl1
45	110	8.4	1701	4	US-10-174-794-5	Sequence 5, Appl1

QY	1	ATGGGTCAAGAGCTGTGTCAGAAAGCTGACAGCTGTGATGATGAGCTGTGACATTTTCA	60
DB	1	ATGGGTCAAGAGCTGTGTCAGAAAGCTGACAGCTGTGATGATGAGCTGTGACATTTTCA	60
QY	61	GAGGAGGCGAGGACACAGCTGTGCGAGAGAGTCAAGCTGAGACCGAGGAGCTGCTTC	120
DB	61	GAGGAGGCGAGGACACAGCTGTGCGAGAGTCAAGCTGAGACCGAGGAGCTGCTTC	120
QY	121	AACTTAACAGCTGTGTCAGAAAGCTGATGCTTCACTTTCACCCAGGAGCTT	180
DB	121	AACTTAACAGCTGTGTCAGAAAGCTGATGCTTCACTTTCACCCAGGAGCTT	180
QY	181	CAAGTCCCGCCGCGCCAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGT	240
DB	181	CAAGTCCCGCCGCGCCAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGT	240
QY	241	CAGGAAGAGATCTGAGGACAGGCTTTTCACTGAGAGAGAGAGAGAGAGAGAGAGAG	300
DB	241	CAGGAAGAGATCTGAGGACAGGCTTTTCACTGAGAGAGAGAGAGAGAGAGAGAGAG	300
QY	301	TCATCTTAATTGAATCTGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG	360
DB	301	TCATCTTAATTGAATCTGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG	360
QY	361	ATTACAGAAATTAATGAGCACTAGTGTGCTTTAAAGTCAATGAGTAAATGAGAGAA	420
DB	361	ATTACAGAAATTAATGAGCACTAGTGTGCTTTAAAGTCAATGAGTAAATGAGAGAA	420

Db	361	ATTAGAGGAATTAATGACAACTGAGGCTTTAAAGTCATCAGCANTGAATGACAGAGAA	420
QY	421	GGAATCCCAATTTACGCTATCCGAAAGCTTCTCTCGTGAAGGGTTGAAACATGCCAAT	480
Db	421	GGAATCCCAATTTACGCTATCCGAAAGCTTCTCTCGTGAAGGGTTGAAACATGCCAAT	480
QY	481	ATTGTGCTCCGACATGACATTAATCCACCAAGAACACATGACATTCGTTTTGAATAC	540
Db	481	ATTGTGCTCCGACATGACATTAATCCACCAAGAACACATGACATTCGTTTTGAATAC	540
QY	541	ATGACACACAGACTGAGCCAGTATATGTCTCAGATCCAGAGAGGCTTCATCTCATAT	600
Db	541	ATGACACACAGACTGAGCCAGTATATGTCTCAGATCCAGAGAGGCTTCATCTCATAT	600
QY	601	GTCAGACTTTTCATGTTTCAACTTTTGCGGGGCTCGCGTACATCCACCAACAGT	660
Db	601	GTCAGACTTTTCATGTTTCAACTTTTGCGGGGCTCGCGTACATCCACCAACAGT	660
QY	661	CTTACAGGGGACCTGAAACCTCAAACTTACATACAGTCACTGGAGAGCTAAACTG	720
Db	661	CTTACAGGGGACCTGAAACCTCAAACTTACATACAGTCACTGGAGAGCTAAACTG	720
QY	721	GCTATTTTGGTCTTGCCCGGGGCAAGTCATTTCCAGCCAGACATTCCTTCAGAAATC	780
Db	721	GCTATTTTGGTCTTGCCCGGGGCAAGTCATTTCCAGCCAGACATTCCTTCAGAAATC	780
QY	781	GTCACCTCTGAGTACCGGCCCCCTGATGCTTTGCTGGAGGCACTGAATATCTCTGAG	840
Db	781	GTCACCTCTGAGTACCGGCCCCCTGATGCTTTGCTGGAGGCACTGAATATCTCTGAG	840
QY	841	CTGACATATATGGGGTGACGGCTGATCTTATTTGAATGTTCCAGGGTCAACCTTTGTT	900
Db	841	CTGACATATATGGGGTGACGGCTGATCTTATTTGAATGTTCCAGGGTCAACCTTTGTT	900
QY	901	CCTGGGGTTTCCAACTCTCTTGAACAGCTGAGAGAAATCTGGAGAGTCTGGAGTCCCT	960
Db	901	CCTGGGGTTTCCAACTCTCTTGAACAGCTGAGAGAAATCTGGAGAGTCTGGAGTCCCT	960
QY	961	ACAGAGGATATCTTGCGCGGAGTCTCCAGCTCAATCAATCCAGAAATGTTCCCA	1020
Db	961	ACAGAGGATATCTTGCGCGGAGTCTCCAGCTCAATCAATCCAGAAATGTTCCCA	1020
QY	1021	CTGCTTACGCTCGAAGCTTCATGTTGTCTGGAACAGGCTGGGCAAGGTTCTGAAAGCT	1080
Db	1021	CTGCTTACGCTCGAAGCTTCATGTTGTCTGGAACAGGCTGGGCAAGGTTCTGAAAGCT	1080
QY	1081	GAAAGACCTGGGCTCCAGATGCTTAAAGGCTTTCCAGAGACCGGAGTCTCCGCCAGGAA	1140
Db	1081	GAAAGACCTGGGCTCCAGATGCTTAAAGGCTTTCCAGAGACCGGAGTCTCCGCCAGGAA	1140
QY	1141	GCACCTGTTCAATGATATTTTCAGGCGCCTGCGCATCTCAGCTGTACACAGTTTCTGATGAG	1200
Db	1141	GCACCTGTTCAATGATATTTTCAGGCGCCTGCGCATCTCAGCTGTACACAGTTTCTGATGAG	1200
QY	1201	GAGTCTTTGTTTACGTTTCAGAGTGAAGGCTTAAAGCCAGAAATGTGTGACCTTTTGCC	1260
Db	1201	GAGTCTTTGTTTACGTTTCAGAGTGAAGGCTTAAAGCCAGAAATGTGTGACCTTTTGCC	1260
QY	1261	TTCCTACAGAAAGGTCAACCCAGCCAGTCTTGGCAAAATGCTGTGA 1308	
Db	1261	TTCCTACAGAAAGGTCAACCCAGCCAGTCTTGGCAAAATGCTGTGA 1308	

```

RESULT 2
US-10-224-562-4
; Sequence 4, Application US/10224562
; Patent No. 6730506
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01098DIV

```

CURRENT APPLICATION NUMBER: US/10/224,562									
CURRENT FILING DATE: 2002-08-21									
NUMBER OF SEQ. ID NOS: 10									
SOFTWARE: FastSeq for Windows Version 4.0									
SEQ. ID NO 4									
LENGTH: 1308									
TYPE: DNA									
ORGANISM: Human									
US-10-224-562-4									
Query Match									
Best Local Similarity 100.0%; Score 1308; DB 4; Length 1308;									
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATGGGTAAAGAGCTGTGTGCMAAAGACTGTACAGCTGGAGTGCACCTGTACATTGTCTCA	60						
DB	1	ATGGGTTCMAAGAGCTGTGTGCMAAAGACTGTACAGCTGGAGTGCACCTGTACATTGTCTCA	60						
QY	61	GAGGAGAGCGAGGACACAGCTGTGCGAGAGTACAGCTGAGACCAAGAGGCTGCCTTC	120						
DB	61	GAGGAGAGCGAGGACACAGCTGTGCGAGAGTACAGCTGAGACCAAGAGGCTGCCTTC	120						
QY	121	AAGCTAACAGACCTTAAAGAAAGCATCATGTTCATGACCTTCATTTCACCCAGGGGACTT	180						
DB	121	AAGCTAACAGACCTTAAAGAAAGCATCATGTTCATGACCTTCATTTCACCCAGGGGACTT	180						
QY	181	CAAGCTGCGCGTGCCGAGAAAGTTCAAGAGTAAAGGCCACGAGTAACAGTATTTGTTT	240						
DB	181	CAAGCTGCGCGTGCCGAGAAAGTTCAAGAGTAAAGGCCACGAGTAACAGTATTTGTTT	240						
QY	241	CAGGAAGAGATCTGAGGACGAGGTTTCAGTGGAGAAAGAGCTCCCTTTTGGGGCAGCC	300						
DB	241	CAGGAAGAGATCTGAGGACGAGGTTTCAGTGGAGAAAGAGCTCCCTTTTGGGGCAGCC	300						
QY	301	TCATCTTACTTGAACCTTGAAGAACTGGTGAAGGCTTTATGCGACAGTTTCAAGGGG	360						
DB	301	TCATCTTACTTGAACCTTGAAGAACTGGTGAAGGCTTTATGCGACAGTTTCAAGGGG	360						
QY	361	ATTAGCAGAAATAATGAGCACTATGTGGCTTTAAAGTCATCAGCATGAAATGACAGGAA	420						
DB	361	ATTAGCAGAAATAATGAGCACTATGTGGCTTTAAAGTCATCAGCATGAAATGACAGGAA	420						
QY	421	GGAATCCATTTCAGGATATCCGAGAGGCTCTCTCCGTAAGGGGTTGAAACATGSCAAT	480						
DB	421	GGAATCCATTTCAGGATATCCGAGAGGCTCTCTCCGTAAGGGGTTGAAACATGSCAAT	480						
QY	481	ATTGTGCTCTGTGCATATATCCACACCAAGAGCACTGACATTCGTTTTGAAATAC	540						
DB	481	ATTGTGCTCTGTGCATATATCCACACCAAGAGCACTGACATTCGTTTTGAAATAC	540						
QY	541	ATGCAACAAGACCTGGCCCCAGTATATGTCTGCAGATCCAGAGGGCTTCATCTCATAT	600						
DB	541	ATGCAACAAGACCTGGCCCCAGTATATGTCTGCAGATCCAGAGGGCTTCATCTCATAT	600						
QY	601	GTGAGACTTTTTCAGTTTCACTTTTGGGGGGCGGGGTATCATCCACCAACAGGTT	660						
DB	601	GTGAGACTTTTTCAGTTTCACTTTTGGGGGGCGGGGTATCATCCACCAACAGGTT	660						
QY	661	CTTACAGGGGACTGGAAGCTCAGAACTTATCTCATAGTCACTTGGAGAGCTCAAACTG	720						
DB	661	CTTACAGGGGACTGGAAGCTCAGAACTTATCTCATAGTCACTTGGAGAGCTCAAACTG	720						
QY	721	GCTGATTTTGGTCTTGCCCGGGGCAAGTCAATCCACGACGACATACCTTTCAAGAGTC	780						
DB	721	GCTGATTTTGGTCTTGCCCGGGGCAAGTCAATCCACGACGACATACCTTTCAAGAGTC	780						
QY	781	GTGACCTCTGTGTACCGGCCCCCTGATGTCTTGTCTGGAGGCACTGAATATTCCTTAAG	840						
DB	781	GTGACCTCTGTGTACCGGCCCCCTGATGTCTTGTCTGGAGGCACTGAATATTCCTTAAG	840						
QY	841	CTGCAACATAGGGGTGAGGCTGCATCTTTATGAATGTTCCAGAGGCTCAACTTTGTTT	900						
DB	841	CTGCAACATAGGGGTGAGGCTGCATCTTTATGAATGTTCCAGAGGCTCAACTTTGTTT	900						

QY	30	TCATCTTA	CTTGAAC	CTTGAGAA	AGCTGGG	TGAAAG	CGCTTTA	TGGA	CAGTTTAA	CAAGGG	360				
Db	418	TCATCTTA	CTTGAAC	CTTGAGAA	AGCTGGG	TGAAAG	CGCTTTA	TGGA	CAGTTTAA	CAAGGG	477				
QY	361	ATTAGCA	GAATAA	ATGACA	CACTAGT	GGCTTTAA	AAATCAT	CAGCAT	GAATGAC	AGGAA	420				
Db	478	ATTAGCA	GAATAA	ATGACA	CACTAGT	GGCTTTAA	AAATCAT	CAGCAT	GAATGAC	AGGAA	537				
QY	421	GGAGT	CCCATTTA	CAGCTAT	CCGAA	AGCTTCTCT	CCTGA	AGGCTTTG	AAACAT	GGCAAT	480				
Db	538	GGAGT	CCCATTTA	CAGCTAT	CCGAA	AGCTTCTCT	CCTGA	AGGCTTTG	AAACAT	GGCAAT	597				
QY	481	ATTGTG	CTCCGCA	TGACAT	TAATCAC	ACCAAGAA	GAC	CTGAC	ATTGCTTTG	GAATAC	540				
Db	598	ATTGTG	CTCCGCA	TGACAT	TAATCAC	ACCAAGAA	GAC	CTGAC	ATTGCTTTG	GAATAC	657				
QY	541	ATGCA	CACAGA	CGTGG	CCAGATAT	ATATGCTC	GAGCAT	CAGAG	GGGCTTCAT	CTCATAT	600				
Db	658	ATGCA	CACAGA	CGTGG	CCAGATAT	ATATGCTC	GAGCAT	CAGAG	GGGCTTCAT	CTCATAT	717				
QY	601	GTCAG	ACTTTTCA	TGTTTCA	CTTTG	GGGGG	CCTGG	CGTAC	TCCAC	CAACCA	CAGTT	660			
Db	718	GTCAG	ACTTTTCA	TGTTTCA	CTTTG	GGGGG	CCTGG	CGTAC	TCCAC	CAACCA	CAGTT	777			
QY	661	CTTAC	AGGGG	CCGAA	AACCTC	GAACTT	TA	CTC	ATCAG	TAC	CCCTGG	AGAGCT	CAAA	CTG	720
Db	778	CTTAC	AGGGG	CCGAA	AACCTC	GAACTT	TA	CTC	ATCAG	TAC	CCCTGG	AGAGCT	CAAA	CTG	837

Db	838	GCTGATTTTGTCCTTGGCCCGGCAAGTCCATTCCTCCAGCCAGACATACCTTTGAGAAGTC	897
Qy	781	GTGACCTCTTGATACCGGCCCCCTGATGCTTTGCTGGAGCATGAAATATTCCTTGAG	840
Db	898	GTGACCTCTTGATACCGGCCCCCTGATGCTTTGCTGGAGCATGAAATATTCCTTGAG	957
Qy	841	CTGACATATGAGGATGAGGCTGCATCTTTATTGAATGTTCCAGGGTCAACCTTTGTTT	900
Db	958	CTGACATATGAGGATGAGGCTGCATCTTTATTGAATGTTCCAGGGTCAACCTTTGTTT	1017
Qy	901	CCTGGGGTTTCCATCTCTTGAACAGCTGGAGAAAATCTGGGAGGTGCTGGAGTCCCT	960
Db	1018	CCTGGGGTTTCCATCTCTTGAACAGCTGGAGAAAATCTGGGAGGTGCTGGAGTCCCT	1077
Qy	961	ACAGAGGATCTTGAGCCGGAGTCTCCAGGTACTTAACAATCCAG	1009
Db	1078	ACAGAGGATCTTGAGCCGGAGTCTCCAGGTACTTAACAATCCAG	1126

```

RESULT 4
US-10-224-562-1
; Sequence 1, Application US/10224562
; Patent No. 6730506
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001098DIV
; CURRENT APPLICATION NUMBER: US/10/224,562
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-224-562-1

Query Match      77.1%;   Score 1009;   DB 4;   Length 2203;
Best Local Similarity 100.0%;   Pctd. No. 0;
Matches 1009;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```

```
QY      1 ATGGGTCAAGAGCTGTGTGCAAGACTGTACAGCTTGATGCACTGCTCAATGTTCA 60
DB      118 ATGGGTCAAGAGCTGTGTGCAAGACTGTACAGCTTGATGCACTGCTCAATGTTCA 177
QY      61 GAGGAGGCGAGGACACAGCTGTGCGAGGAGTCAAGCTGACCAACGAGGCGCTTC 120
DB      178 GAGGAGGCGAGGACACAGCTGTGCGAGGAGTCAAGCTGACCAACGAGGCGCTTC 237
QY      121 AAGCTAACAGACTTAAAGAGCATCATGTTCCATGCTTCAATTCACCCAGGGAATT 180
DB      238 AAGCTAACAGACTTAAAGAGCATCATGTTCCATGCTTCAATTCACCCAGGGAATT 297
QY      181 CAAGCTGCGCGTCCCGAAGTTCAGAGTAAAGGCAAGAGTAAAGATGTTT 240
DB      298 CAAGCTGCGCGTCCCGAAGTTCAGAGTAAAGGCAAGAGTAAAGATGTTT 357
QY      241 CAGGAAGAGATCTGAGGAGGCTTTCAGTGAAGAGAGCCCTCTTTGGGGGAGCC 300
DB      358 CAGGAAGAGATCTGAGGAGGCTTTCAGTGAAGAGAGCCCTCTTTGGGGGAGCC 417
QY      301 TCATCTTACTTGAACCTTGAGAAAGCTGGTGAAGCTTTATGCAAGTTTCAAGGG 360
DB      418 TCATCTTACTTGAACCTTGAGAAAGCTGGTGAAGCTTTATGCAAGTTTCAAGGG 477
QY      361 ATTAGAGATTAATGAGCAACTAGTGTCTTAAAGTCAATGCAATGAGAGAA 420
DB      478 ATTAGAGATTAATGAGCAACTAGTGTCTTAAAGTCAATGAGAGAA 537
QY      421 GAGTCCCATTTACAGCTATCCGAGAGCTTCTCTGTAAGGGTTGAAACATGCAAT 480
DB      538 GAGTCCCATTTACAGCTATCCGAGAGCTTCTCTGTAAGGGTTGAAACATGCAAT 597
QY      481 ATTGTGCTCTGATGATGATATTCACACCAAGAGCACTGACATTCGTTTGAATAC 540
DB      598 ATTGTGCTCTGATGATGATATTCACACCAAGAGCACTGACATTCGTTTGAATAC 657
QY      541 ATGCACACAGACCTGGGCCAGTATATGTCTCAGCATCCAGAGGGCTTCAATCTCAAT 600
DB      658 ATGCACACAGACCTGGGCCAGTATATGTCTCAGCATCCAGAGGGCTTCAATCTCAAT 717
QY      601 GTGAGACTTTTCAATGTTTCACTTTTGCAGGAGCTGAGTCAATCCACCAACAGCTT 660
DB      718 GTGAGACTTTTCAATGTTTCACTTTTGCAGGAGCTGAGTCAATCCACCAACAGCTT 777
QY      661 CTTCAAGAGGACCTGAAACCTGAACTTATGATGATGATGATGATGATGATGATG 720
DB      778 CTTCAAGAGGACCTGAAACCTGAACTTATGATGATGATGATGATGATGATGATG 837
QY      721 GCTGATTTTGTCTGTCGCGGAGGCAAGTCAATCCAGAGCAAGATCTTCAAGAGTC 780
DB      838 GCTGATTTTGTCTGTCGCGGAGGCAAGTCAATCCAGAGCAAGATCTTCAAGAGTC 897
QY      781 GTGACCTCTGTGATCCGGGCCCTGTATGCTTGTGCTGGAGCACTGAAATATCTCTGAG 840
DB      898 GTGACCTCTGTGATCCGGGCCCTGTATGCTTGTGCTGGAGCACTGAAATATCTCTGAG 957
QY      841 CTGGAATATGAGGCTGAGGCTGATCTTATTTAATGTTTCAAGGCTCAACTTTGTTT 900
DB      958 CTGGAATATGAGGCTGAGGCTGATCTTATTTAATGTTTCAAGGCTCAACTTTGTTT 1017
QY      901 CCGGGGTTTCCAACTCTTGAAGAGCTGAGAAATCTGGGAGGCTGCGGAGTCCCT 960
DB      1018 CCGGGGTTTCCAACTCTTGAAGAGCTGAGAAATCTGGGAGGCTGCGGAGTCCCT 1077
QY      961 ACAGAGGATATCTTGGCGGAGTCTTCAAGCTAATCAATCAATCCAG 1009
DB      1078 ACAGAGGATATCTTGGCGGAGTCTTCAAGCTAATCAATCAATCCAG 1126
```

RESULT 5
US-09-387-212-14
; Sequence 14, Application US/09387212A

```
; Patent No. 6309849
; GENERAL INFORMATION:
; APPLICANT: ROBINSON, KEITH E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; FILE REFERENCE: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/09/387, 212A
; CURRENT FILING DATE: 1999-08-31
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 14
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-387-212-14

Query Match      36.8%; Score 481; DB 3; Length 481;
Best Local Similarity 100.0%; Pred. No. 2.6e-149; Mismatches 0; Gaps 0;
Matches 481; Conservative 0; Indels 0;

QY      720 GCGTATTTTGTGTCCTTGCCTGGGCAAGTCAATCCAGCCAGACATCTTTCAAGAGT 779
DB      1 GCGTATTTTGTGTCCTTGCCTGGGCAAGTCAATCCAGCCAGACATCTTTCAAGAGT 60
QY      780 CCGTACCCTCTGTACCGGCCCCCTGATGCTTTGCTGGAGGCACTGAATATTCCTTGA 839
DB      61 CCGTACCCTCTGTACCGGCCCCCTGATGCTTTGCTGGAGGCACTGAATATTCCTTGA 120
QY      840 GCTGACATATGAGGCTGAGGCTGATCTTATGAAATGTTCCAGGGTCAACTTTGTT 899
DB      121 GCTGACATATGAGGCTGAGGCTGATCTTATGAAATGTTCCAGGGTCAACTTTGTT 180
QY      900 TCTGTGGGTTTCCAACTCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCC 959
DB      181 TCTGTGGGTTTCCAACTCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCC 240
QY      960 TACAGAGATTAATGAGGCTGAGGAGTCTCAAGTACTCAATCAATCAAGATGTTTCC 1019
DB      241 TACAGAGATTAATGAGGCTGAGGAGTCTCAAGTACTCAATCAATCAAGATGTTTCC 300
QY      1020 ACTGCTTACGCTCCAGACCTTATGTTGCTGAAACAGGCTGAGGCTTCTGAAGC 1079
DB      301 ACTGCTTACGCTCCAGACCTTATGTTGCTGAAACAGGCTGAGGCTTCTGAAGC 360
QY      1080 TGAAGACTGCGCTCTCCAGATGCTTAAAGGCTTTCCAGAGACCGGCTTCCGACAGA 1139
DB      361 TGAAGACTGCGCTCTCCAGATGCTTAAAGGCTTTCCAGAGACCGGCTTCCGACAGA 420
QY      1140 AGCACTTGTATGATTAATTTTCAAGGCGCTGCAATGAGCTGTACAGTCTTCGATGA 1199
DB      421 AGCACTTGTATGATTAATTTTCAAGGCGCTGCAATGAGCTGTACAGTCTTCGATGA 480
QY      1200 G 1200
DB      481 G 481

RESULT 6
US-09-948-802-14
; Sequence 14, Application US/09948802
; Patent No. 6465232
; GENERAL INFORMATION:
; APPLICANT: ROBINSON, KEITH E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; FILE REFERENCE: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/09/948, 802
; CURRENT FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/387, 212
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 14
```

LENGTH: 481
TYPE: DNA
ORGANISM: Homo sapiens
US-09-948-802-14

Query Match 36.8%; Score 481; DB 4; Length 481;
Best Local Similarity 100.0%; Pred. No. 2.6e-149;
Matches 481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 720 GGCTGATTTTGGCTTTCGCCGCGGCAAGTCCATTTCCAGCAGACATCTTCCAGAAAT
DB 1 GGCTGATTTTGGCTTTCGCCGCGGCAAGTCCATTTCCAGCAGACATCTTCCAGAAAT
QY 780 CGTGACCTCTGCTGACCGGCGCCCTGATGCTTGGCGAGCACTGAATATTTCTCTGA
DB 61 CGTGACCTCTGCTGACCGGCGCCCTGATGCTTGGCGAGCCATGAATATTTCTCTGA
QY 840 GCTGACATATGGGGTGGCAGGCTGCATCTTATTTGAATGTTCCAGGGTCAACTTTGTT
DB 121 GCTGACATATGGGGTGGCAGGCTGCATCTTATTTGAATGTTCCAGGGTCAACTTTGTT
QY 900 TCCCTGGGGTTTCCACATCTTTCAGAGCTGAGAAATCTGGAGAGTGTGGAGTCCC
DB 181 TCCCTGGGGTTTCCACATCTTTCAGAGCTGAGAAATCTGGAGAGTGTGGAGTCCC
QY 960 TACAGAGATCTTGGCGCGGAGTCTTCAGAGTACCTAATCAATCCAGAAATGTTTCCC
DB 241 TACAGAGATCTTGGCGCGGAGTCTTCAGAGTACCTAATCAATCCAGAAATGTTTCCC
QY 1020 ACTGACCTAGGCGCTCGAAGCCTTCATGTTCTGGAACAGGCTGGGAGGTTCTGTAAGC
DB 301 ACTGACCTAGGCGCTCGAAGCCTTCATGTTCTGGAACAGGCTGGGAGGTTCTGTAAGC
QY 1080 TGAAGACTGCTGCCCTCCAGATGCTAAAGGCTTTCACAGAGACCGGCTCTCCGCGAGGA
DB 361 TGAAGACTGCTGCCCTCCAGATGCTAAAGGCTTTCACAGAGACCGGCTCTCCGCGAGGA
QY 1140 AGCAGCTTGTTCATGATTTATTTACGCGCCCTGCAATCTCAGCTGTACAGCTTCTGATGA
DB 421 AGCAGCTTGTTCATGATTTATTTACGCGCCCTGCAATCTCAGCTGTACAGCTTCTGATGA
QY 1200 G 1200
DB 481 G 481

RESULT 7

US-09-206-344A-3
Sequence 3, Application US/09206344A
Patent No. 6432668
GENERAL INFORMATION:
APPLICANT: Christoph Reinhard
APPLICANT: David Pot
APPLICANT: Aliaf Kassam
APPLICANT: Tasha Marenbach
APPLICANT: Lewis T. Williams
TITLE OF INVENTION: HUMAN CYCLIN-DEPENDENT KINASE (hPFTAIRB)
FILE REFERENCE: 200103.445 / 1429.002
CURRENT APPLICATION NUMBER: US/09/206,344A
CURRENT FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 2467
TYPE: DNA
ORGANISM: mouse
US-09-206-344A-3

Query Match 33.5%; Score 438; DB 4; Length 2467;
Best Local Similarity 65.8%; Pred. No. 1.5e-134;
Matches 636; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

QY 277 AAGAGCTCCCTTTGGGCGAGGCTCATCTTACTTGAACCTTGAGAAAGCTGGGTGAAGGC 336

DB 353 ACGAGTCCCAAAATTTGGAAAAAGCTGACTCATACGAAAACTGAAAAAATCTGGGGAGGA 412
QY 337 TCTTATGCGACAGATTATACAAAGGAGATTACAGAAATTAATGACAACTAGTGGCTTTAAAA 396
DB 413 TCTTATGCGACAGATTATACAAAGGAGAAAAAGCAAGTAAATGGGAAGCTGTGGCTTTAAAG 472
QY 397 GTCATGACATGATATGACAGAGAGAGAGATCCATTATACGCTATCCGAGAACTTTCTTC 456
DB 473 GTGATCCGCTCAGAGAAAGAGAGGACACCTTTTACAGCCATCGGGAAGGTTCCCTG 532
QY 457 CTGAAGGTTTGAACATGCCAATATTGTCCTCGGATGACATATATCCACCAAGAG 516
DB 533 TTGAAGAGACTAAGACCCCAACATCGTTGCTTACAGACATATCAACTTAAGAA 592
QY 517 ACACTGACATTTGCTTTTGAATATACACAGACAGACCTGCGCCAGTATATGTTCTCAGAT 576
DB 593 ACCCTGACCTTGTCTTTTAATATAGTGACACATGATTTATGTCAATGACAAAGCAC 652
QY 577 CCAGAGGGCTTCATCTCATATATGTCAGACTTTTCATGTTTCACTTTTGGGGGCTTG 636
DB 653 CCGAGAGGACTTCATTCAGATATATGTAAGTTTATTTTACGCTGCTGGAGAGACTG 712
QY 637 GGGTACATGACACCAACCAACGCTTCTCAAGGAGCTGAAACCTCAGAACTTACTCATC 696
DB 713 TCTTACATCCACGAGCTTATATTTTGAACAGACCTGAAACCGCAGAACTTCTCATC 772
QY 697 ACTCACCTGGAGAGCTCAAACTGCTGATTTTGTCTTGGCCGCGGCAAGTCCATATCCC 756
DB 773 ACGATAGGGGGAGATGTAAGTGGAGATTTCCGCTCTGGCAAGAACAAATCCGCTCC 832
QY 757 ACGCAGACATCTTCTCAGAGTCTGACCTCTGCTGATCCGCGCCCTGATGCTTGTCTG 816
DB 833 ACGCAGACATCTTCTCAGAGTCTGACCTCTGCTGATCCGCGCCCTGATGCTTGTCTG 892
QY 817 GGAAGCAGTGAATATTTCCCTGAGCTGACATATGGGGGTCAGGCTGCATCTTATTTGAA 876
DB 893 GGCTTACAGAAATTTTCCACCTGCTTACATGTTGGAGGTTGGCTGTATCTTCTTGAG 952
QY 877 ATGTTCCAGGCTCAACCTTTGTTTCTGCGGCTTTCACATCTTGAACAGCTGAGAAA 936
DB 953 ATGATCCAAAGAGTGTGCTGCTGTTCCAGGAATGAAAAGATCAAGATCACTTGAACGG 1012
QY 937 ATCTGGAGGCTGCTGGAGTCCCTACAGAGATCTTGGCCGAGAGTCTTCAAGCTACT 996
DB 1013 ATATTTCTGTTCTTGGAAACACGATGAGGACAGGTGCTGAGTTCTATTTTACCA 1072
QY 997 AACTCAATCCAGAAATGTTTCCACCTGCTAGCCCTCGAAGGCTTCATGTTCTGGAAC 1056
DB 1073 CATTTTAAGCCAGAAAGCTTTTACCGGTACCACTTAAAGCTTACAGACATGGAAT 1132
QY 1057 AGGCTGGGAGAGGTTCTGTAAGCTGAAGACCTGCTCCAGATGCTAAAGGCTTTTCCC 1116
DB 1133 AAGCTAGCTATGTAATATATCTGAAGCTTGGCTTCCAACTTCTCCAGTGTCCCA 1192
QY 1117 AAGAACCGGCTCTCGGCCCAAGACATTTGTCATGATTTATTTAGCGCCCTGCCATCT 1176
DB 1193 AAGAACAGGCTATACAGACAGGCGCCCTTGAACCATGATATTTAGCATCTGCTCCA 1252
QY 1177 CAGCTGTACAGCTTCCGATGAGAGAGTCTTTTATACGTTTACAGAGTGAAGGCTTAA 1236
DB 1253 CGGCTATGGAGCTGATGATATGTTCTTATTTTACCGTCCCAAAATGATGATGCAA 1312
QY 1237 CCAGAA 1242
DB 1313 CCAGAA 1318

RESULT 8

US-09-206-344A-1
Sequence 1, Application US/09206344A
Patent No. 6432668
GENERAL INFORMATION:

```

; APPLICANT: Christoph Reinhard
; APPLICANT: David Pot
; APPLICANT: Alfat Kassem
; APPLICANT: Tasba Marendach
; APPLICANT: Lewis T. Williams
; TITLE OF INVENTION: HUMAN CYCLIN-DEPENDENT KINASE (hPPTAIRE)
; FILE REFERENCE: 200103.445 / 1429.002
; CURRENT APPLICATION NUMBER: US/09/206.344A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1272
; TYPE: DNA
; ORGANISM: human
US-09-206-344A-1
```

```

Query Match      33.5%; Score 437.8; DB 4; Length 1272;
Best Local Similarity 65.1%; Pred. No. 1.1e-134;
Matches 646; Conservative 0; Mismatches 347; Indels 0; Gaps 0;
```

```

QY 280 AGCTCCCTTTGGGCGAGCTCATCTTACTTGAAGTGGAGAGCTGGGTGAAGCTCT 339
DB 238 AGTCCCAATTTGGAAAAGCTGACTATGAAAAGCTGAGAAAAGCTAGGGAGGATCT 237
QY 340 TATGCCAGATTTCAGAGGATTTAGAGAAATTAATGGAACAACCTAGTGGCTTTAAAGTC 339
DB 298 TATGCTACAGATATACAAAGGAAAAGGAGGTAATGGAAAGTTGGTGAAGCTGAAAGTG 357
QY 400 ATCAGCATGATCAGAGGAGAGAGTCCATTACAGCTATCCGAGAAGCTTCTCTCTG 459
DB 358 ATCAGGCTGAGAGAAAGAGAGGAGACCTTTACAGCTATCAGGAGGCTTCTTTTA 417
QY 460 AAGGTTTGAACAATGCCAATATTGTGCTCTTGATGACATTAATCCAACCAAGAGACA 519
DB 418 AAGAGCTAAACATGCTAATCATATGCTACTTCATGACATCATCAACCAAGAGACG 477
QY 520 CTGACATTCGTTTTGAATATCATGACACAGACCTGGCCCAAGATATGTCTCAGCATCCA 579
DB 478 CTGACATTCGTTTTGAATATGTGACACATGATTTATGTACATGACAGACACCTCT 537
QY 580 GAGGGCTTCATCTCATTAATGTCAAGCTTTTCATGTTTCACTTTTGGCGGGCTGGCG 639
DB 538 GGGGGCTGATCCAGATATATGGAAGTGTTTTATTGAGTTGTGCGAAGGTCTGTCT 597
QY 640 TACATCCACACCAACACGCTTTTCACAGGAGCTGAAACCTAGAACTTATCTCATAGT 699
DB 598 TACATCCACACGCTTATATTTTGCACAGAGACCTGAAACCAAGAACCTTCTGATCAGT 657
QY 700 CACCTGGAGAGCTCAAACTGCTGATTTTGTCTTGGCCCGGGCCCAAGTCCATTCACAG 759
DB 658 GACACGGGGAGTTAAAGCTGGCAATTTGCTTGGCAAGACAAATCCGTCCTAGG 717
QY 760 CAGACATCTCTTCAGAAAGTGTGACCTCTGTGTAACGGCCCTCATGATCTTGTGCGGA 819
DB 718 CACACATCTCTCCAAAGAGTGTGTAACCTGTGTAACAGCTTCAATGTCTTACAGG 777
QY 820 GGCATCGAATATCTCTGAGCTGACATATGGGGTGCAGGCTGCATCTTTATTTGAATG 879
DB 778 TCAACGAATATTCACCTGCTTGACATGTGGGGAGTGGTTCATCTTTGTTGAAAG 837
QY 880 TTCCAGGGTCAACCTTTGTTTCTGGGGTTTCCAAATCCTTGAACAGCTGGAGAAATC 939
DB 838 ATCCAGAGAGTCTGCTTTTCCAGAAATGAAGAAATTCAGATTAACCTTGAAGAAAT 897
QY 940 TGGAGAGTGTGGAGATCTCTTCAAGAGATTAATTTGGCCGGAGTCTTCAAGTCACTTAAC 999
DB 898 TTTCTGTTCTTGGACACCAAAATGAGGACATGCTGAGGTTCAATCTTTACACAGT 957
QY 1000 TACAATCCAGAAATGTGTCCCATGCTTACGCTTCGAAGCTTCAATTTGTCTGAAACAG 1059
DB 958 TTTAAGCCAGAAAGCTTTACCTGTACAGCTTAAAAACCTTAGACAGACATGAGATAG 1017
```

```

QY 1060 CTGGCGAGGGTCTCTGAAGCTGAAGACCTGGCTCTCCAGATGCTAAAGCTTTCCAGA 1119
DB 1018 CTCAGCTATGTGAACCATGACAGAGGACCTGGCTCCAGACTCTCAATGTTCCCAAG 1077
QY 1120 GACCGGCTCCGCGCCAGAGAGCACTGTTCAATGATTAATTTACAGGGCCCTGCATCTCAG 1179
DB 1078 AACAGCTGTCCGACAGGCTGCTTGAAGCCAGCATATTTATGTACCTGCGCCACAG 1137
QY 1180 CTGTACAGCTTCTCTGATGAGAGTCTTTGTTTACAGTTTACAGATGAGGCTAAAGCCA 1239
DB 1138 CTATGGAACTACCGACAGATGCTTCTATTTTATCTGTCCCAAGATGTGATGACACA 1197
QY 1240 GAAATGTGACCTTTTGGCTCTTACCGAAT 1272
DB 1198 GAAAGTGAAGAAAGCATGCGGCTTTGGGAAA 1230
```

RESULT 9

```

US-09-206-344A-13
; Sequence 13, Application US/09206344A
; Patent No. 643268
; GENERAL INFORMATION:
```

```

; APPLICANT: Christoph Reinhard
; APPLICANT: David Pot
```

```

; APPLICANT: Alfat Kassem
; APPLICANT: Tasba Marendach
```

```

; APPLICANT: Lewis T. Williams
; TITLE OF INVENTION: HUMAN CYCLIN-DEPENDENT KINASE (hPPTAIRE)
```

```

; FILE REFERENCE: 200103.445 / 1429.002
; CURRENT APPLICATION NUMBER: US/09/206.344A
```

```

; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 13
```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
```

```

; LENGTH: 1549
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
US-09-206-344A-13
```

```

Query Match      33.3%; Score 436.2; DB 4; Length 1549;
Best Local Similarity 65.0%; Pred. No. 4.3e-134;
Matches 645; Conservative 0; Mismatches 348; Indels 0; Gaps 0;
```

```

QY 280 AGCTCCCTTTGGGCGAGCTCATCTTACTTGAAGTGGAGAGCTGGGTGAAGCTCT 339
DB 451 AGTCCCAATTTGGAAAAGCTGACTCATATGAAAAGCTGGAACCACTAGGGAGAGATCT 510
QY 340 TATGCCAGATTTCAGAGGATTTAGAGAAATTAATGGAACAATGAGGCTTTAAAGTC 399
DB 298 TATGCTACAGATATACAAAGGAAAAGGAGGTAATGGAAAGTTGTAGCTTGAAGTG 570
QY 400 ATCAGCATGATCAGAGGAGAGAGTCCATTTCAGCTATCCAGGAAAGCTTCTCTCTG 459
DB 571 ATCAGGCTGAGAGAGAGAGAGGAGCACTTTCAAGCTATCAGGATCAGGAGCTTCTTTTA 630
QY 460 AAGGTTTGAACAATGCCAATATTGTGCTCTGCAATGACATTAATCCAACCAAGAGACA 519
DB 631 AAGAGCTTAAACATGCTTACATAGTGTACTTCAATGACATTCATCAACCAAGAGAGC 690
QY 520 CTGACATTCGTTTTGAATATCATGACACAGACCTGGCCAGTATATGTCTCAGATCCA 579
DB 691 CTGACATTCGTTGTTGAATATGTGACACCTGATTTATGTCTCAGTACATGAGACAGCCT 750
QY 580 GAGGGCTTCATCTCATTAATGTGACACTTTTCATGTTTCACTTTTGGGGGCTGGCG 639
DB 751 GGGGGGCTGACATCCAGATTAATGTGAAGTGTTTTATTTACGTGCTCGAGGTCTGTCT 810
QY 640 TACAATCCACACCAACAGCTTCTTCAAGGAGACCTGAAACCTGCAATTTACTCATCGT 699
DB 811 TACATTCACACGCTTATTAATTTTGCACAGAGACCTGAAACCAAGAACCTTCTGATCAGT 870
QY 700 CACCTGGAGAGCTCAAACTGCTGATTTTGTCTTGGCCCGGGCCAAAGTCCATTTCCAGC 759
```

```

Db      871 GACAGCGGGGAGTTAAAGCTGCGAGATTTCGGTCTTGCAAGACGAAATCCCTCCTAGC 930
QY      760 CAGACATCTCTTTCAGAACTGCTGACCTCTGTTACCGGCCCCCTGATGCTTTGCTGGGA 819
Db      931 CACACATCTCTCCAAACGAGGTGTTACTTGTGTACAGACCTCCAGATGCTCTTAGGCG 990
QY      820 GCCACGAAATATTCCTCTAGCTGCAATATGGGGGCGAGGCTGCATCTTTATGGAAG 879
Db      991 TCACACGAAATATTCCTCTAGCTGCAATATGGGGGCGAGGCTGCATCTTTGTTGAAATG 1050
QY      880 TTCACGAGTCAACCTTTGTTCTCTGGGGTTTCCAATCTCTGAAACAGCTGAGAAATC 939
Db      1051 ATCCAGAGAGTGTCTCTCTTTTCCAGGAATGAAGACATTCAGATCACTTGACGAAAT 1110
QY      940 TGGGAGCTCTGGGAGTCCCTTCAAGAGATCTTGGCCGGAGTCTCCAGCTACCTAAC 999
Db      1111 TTTCTGTTCTTGGAAACACCAATGAGGACACATGGCTGAGATCATTTCTTACCACTT 1170
QY      1000 TACAATCAGAAATGTTTCCACCTGCTAGCGCTCGAAGCTTCATGTTGTGGAACAG 1059
Db      1171 TTTAAGCCGAAACGCTTACCTGTACAGCTTAAACCTTAGACAGATGGAATPAAG 1230
QY      1060 CTGGGAGGAGTTCCTGAAGCTGAAAGCTGAGCTCCAGATGCTAAAGGCTTTCCAGA 1119
Db      1231 CTCAGCTATGTGAACATGACAGAGACCTGGCTCCAGCTCTCAATATGTTCCCAAG 1290
QY      1120 GACCGGCTCTCGCCGAGAAAGCATTTGATGATTATTCAGCGCCCTGCATCTCAG 1179
Db      1291 AACAGACTGTGGCAGAGCTGCTGAGCAGAGATTTTATGACCTGCGGCGACCG 1350
QY      1180 CTGTACACACTTCTCTATAGAGATCTTTGTTACAGTTTCAGAGTGAAGGCTTAAGCCA 1239
Db      1351 CTATGGAACTCAACCAATGCTCTTCTATTTTAACTGTCCCAATGATGAGATTGCAACCA 1410
QY      1240 GAAATGTGACCTTTTGGCTCTTACAGAAA 1272
Db      1411 GAAAGCTGAGAAAGCATGCGGCTTTGGGAAA 1443

RESULT 10
US-09-620-312D-53
; Sequence 53, Application US/09620312D
; Patent No. 6559662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyun
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonhong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungling
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 53

```

```

; LENGTH: 2868
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (431)..(1924)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2868)
; OTHER INFORMATION: n = a,t,c or g
US-09-620-312D-53

Query Match      22.6%; Score 296.2; DB 4; Length 2868;
Best Local Similarity 58.6%; Pred. No. 2,1e-87;
Matches 555; Conservative 0; Mismatches 383; Indels 9; Gaps 2;

QY      289 TTTGGGCGACCTCATCTTACTTGAACCTGAGAAAGCTGGTGAAGCTCTTATGCGACA 348
Db      905 TTTGGGAACTGGAACCTCATCTTAACTGAGCAAACTGGCGAGGATCTATGCGACC 964
QY      349 GTTTACAAAGGGATTAAGCAATTAATGACAACTAGTGGCTTTTAAAGTCATCAGCATG 408
Db      965 GTCTCAAAAGGCAAAAGCAAGCTCAAGACAACTTGTGGCACTCAAGAGATCAGACTG 1024
QY      409 AATGCAAGAGAGAGTCCCATTTACAGTATCCGAGAAAGCTTCTCTGAAAGGTTTG 468
Db      1025 GAACATGAAAGAGGGGCAACCTGCAACCGCAATCCGGAAGTGTCCCTGCTCAAGAACCTC 1084
QY      469 AAACATGCCAATATTGTGCTCTGATGACATATTCACACCAAGAGACATGACATTTC 528
Db      1085 AAACAGGCCAATATGTTACGTTACATGACATTAATTCACAGAGAGATCCCTCACCTT 1144
QY      529 GTTTTGAATATACATGACACAGACCTGGCCAGTATATGTCTCAGATTCAGAGAGGCTT 588
Db      1145 GTCTTGAAGTACTGGAAGAGACCTGGAAGCAGTACTGGATGACTGTGGGAACATCATC 1204
QY      589 CATCTCATATATGACATTTTCAATGTTTCACTTTTGGGGGCTGGCGTACATCCAC 648
Db      1205 AACATGCAACAGTAACTGTTCTGTTCCAGCTGCTCCGGGCTGGCTCACTGCCAC 1264
QY      649 CACCAACAGTCTTCAAGAGGACCTGAAACCTGAAACCTTACTCATCGTCACTGAGGA 708
Db      1265 CGGCAAGAGTCTCAACCGAGATCTCAAGCCCAAGACCTGCTATCAAGAGAGGGA 1324
QY      709 GAGCTCAACTGCTGATTTTGTCTTGGCCGGGCAAGTCCATTTCCAGCCAGACATAC 768
Db      1325 GAGCTCAAGCTGCTGACTTTGGCTGGCCGAGCAAGTCAATCCAAACAAAGACATAC 1384
QY      769 TCTTCAGAAATGTCGACCTCTGTGTAACGGGCCCTGATGCTTTGCTGGAGCCACTGA 828
Db      1385 TCCAATGAGTGTGACACTGTGTGTAACGGGCCCTGACATCTCTGTGGGTCCACGGAC 1444
QY      829 TATTCCTGAGCTGACATATAGGGGTGAGGCTGACATCTTATGAATGTTCCAGGAT 888
Db      1445 TACTCATCTCAGATTGACATGTGGGATGTGGGCTGCACTTTTATGAGATGGCCACAGGC 1504
QY      889 CAACCTTGTCTCTGAGGATTTCCAAACATCTTGAACAGCTGAGAAATCTGGAGAGTG 948
Db      1505 CGTCCCTCTTCC---GGGCTCCACGGTGAAGAGAAAGCTACATTCATCTTCGGTATC 1561
QY      949 CTGGAGTCCCTACAGAGATATCTGGCCGGAGATCTCCAGACTCATCAATCAATCA 1008
Db      1562 TTAGGAACCCCACTGAGAGAGACGTGGCCAGG-----CATCTGTCCAAACAGAGATTTC 1615
QY      1009 GAATGTTCCCACTGACCTGACGCTGAGAGCTTCAATGTTGTGTGAACAGGCTGGGACAG 1068
Db      1616 AAGACATACACTATCCCAAGTACGAGCCGAGGCTTTTGAAGCAAGACCCGACTT 1675
QY      1069 GTTCTGAAGCTGAAGACCTGGCTCCAGATGCTTAAAGCTTTCCAGAGACCGCGTC 1128
Db      1676 GATAGCGAGGGGCGGACCTCTCTCAACAGCTGTGCAAGGTTGAAGGTCGAAATGAGATC 1735
QY      1129 TCCGCCCAAGAAAGCATTTGTTATGATTTTCAAGCGCCCTGCCATCTCAGCTGATCAG 1188

```

Db 1736 TCCGAGAGAGATGCCATGAAACATCCATCTTCTCAGTCTGGGGAGCGGATCCACAA 1795
Qy 1189 CTTCTGATGAGAGCTTTGTTTACAGTTTCAGAGAGGCTAA 1235
Db 1796 CTTCCGACACTCTTCCATATTGCACTAAGAGATTCAGCTACA 1842

RESULT 11

US-09-016-434-1283
; Sequence 1283, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1283:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g297101
; US-09-016-434-1283

Query Match 22.3%; Score 291.8; DB 4; Length 1242;
Best local Similarity 58.2%; Pred. No. 3,4e-86;
Matches 572; Conservative 0; Mismatches 402; Indels 9; Gaps 3;

Qy 289 TTTGGGCGACCCATCTTCTGAACTTGAGAGAGCTGGTGAAGGCTTTATNGGACA 348
Db 130 TTTGGGAAATGGAACATACGTGAACCTGACAACTGGAGAGGGGACCTATNGCCACA 189
Qy 349 GTTACAGAGGAGATTAGCAGAAATGACAACTAGTGGCTTTAAAGTCATCAGCATG 408
Db 190 GTCTTCAAGGGGCGGACGAACTGACGGAACCTTGTGGCCCTGAAGAGATCCGGCTG 249
Qy 409 AATGACAGAGAGAGTCCCATTTTACAGCTATTCGAGAGAGCTTCTCTCTGAAGGTTTG 468
Db 250 GAGCAGCAGAGAGAGGAGCGCCCTGCACTGCATCCGAGAGGTGTCTGTGTAAGAACCTG 309
Qy 469 AAGCATGCCAATATGTGTCTCTGCAATGATATCCACCAAGAGACACTGACATTC 528

Db 310 AAGCAGCGCAATATTGTATGATCCTTGATATCACTCATCCACAGATGCGCTCCCTCACCGT 369
Qy 529 GTTTTGAATATGACACACAGACCTGGGCCAGATATATGTCTCAGACATCCAGAGAGGCTT 588
Db 370 GTTTTGAATATGACACAGATGATCCTGAAGCAGTATCTGACACATCTGGGAACTCATG 429
Qy 589 CATCTCTAATATGTATGATCTTTTCAATTTTCACTTTTGGGGGGCGTGGCATTCAC 648
Db 430 AGCATGCAACAGTGAATATTTTCATGTTTCAAGCTGTCTCGGGGCTGCGCTACTGTAC 489
Qy 649 CACCAACAGTTCCTTCAAGGAGCTGAAACCTGAACTTCAATCATGATCACTGGGA 708
Db 490 ACCGGAAGATCTGACACCGGAGCTGAAAGCCGAGAACCTGCTATCAAGAGAGGGGG 549
Qy 709 GAGCTCAACCTGCTGATTTTGTCTTCCCGGGGCAAGTCAATTCAGCCAGCATAC 768
Db 550 GAGCTGAAGCTGGCGACCTTTGAGCTGGGCGCAAGGCTCAGTCCCAAGAGCTTAC 609
Qy 769 TCTTCAAGAGTGTGACCTCTGTGACCGGCCCCGTGATGCTTTGCTGGAGCCACTGA 828
Db 610 TCCATGAGGTGTGACCTCTGTGACAGGCCCCGATGTGCTGTGAGATCCAGAG 669
Qy 829 TATTCCTGAGCTGACATATGAGGTGCAAGCTGATCTTTATTTGAATGTTCCAGGAT 888
Db 670 TACTCACCACCCATGCTATGTGGGGCGTGGGCTGATCACTAGAGATGGCCACAGGG 729
Qy 889 CAACCTTGTCTTCTGAGGATTTCCAACTCTTGAACAGCTGAGAGAAATCTGGAGGAT 948
Db 730 AGGCCCTCTTCTCC---GGGCTCCAGTCAAGGAGGAGCTGACCTCATCTTTGCGCTC 786
Qy 949 CTGGAGTCCCTACAGAGATTAATTTGGCGGGAGCTCAAGCTACTAATCAATCA 1008
Db 787 CTGGAGATCCCAAGAGAGAGCTGGCCCGGAGTCAACCGCTTCTGAGTTCGGAC 846
Qy 1009 GAATGTTCCCACTGCTACGCTCGAAGCTTCAATGTTGTGAGAGAGAGCTGGAG 1068
Db 847 TACAGCTTC--CTGTCTACTCTCCGAGCC---GCTCATCAACAGCGCCGAGGTTG 900
Qy 1069 GTTCTGAACTGAAAGACCTGGGCTCCAGATGCTTAAAGGCTTTCCAGAGACCGCTC 1128
Db 901 GATACGAGATGCGATCACCCTCTGAGAGAGCTGCTGTGTATGAATCCAGAGTGCATG 960
Qy 1129 TCCGCGCCAGAGACATTTGATATTTTCAAGGCGCTGCACTCAGCTGATCAG 1188
Db 961 TAGAGAGAGCTGGCTGATGATCTTCTGAGTCTTCCGCTCTGGAGAGCTGTGACCAAG 1020
Qy 1189 CTTCTGATGAGAGCTTTTGTATCAAGTTTCAGAGAGTGAAGCTTAAAGCAGAAATGTGT 1248
Db 1021 CTTGAAGACATGCGCTTCATCTTCCCTGAAGGAGATCCAGCTCCAGAGAGCCAGGC 1080
Qy 1249 GACCTTTTGGCTCTTACCAAA 1271
Db 1081 TACGAGAGCTTGGCTTCCAGCA 1103

RESULT 12

US-08-154-915-1
; Sequence 1, Application US/08154915
; Patent No. 5618669
; GENERAL INFORMATION:
; APPLICANT: Beach, David
; APPLICANT: Xiong, Yue
; TITLE OF INVENTION: Cyclin Complex Rearrangement and Uses
; TITLE OF INVENTION: Related Thereeto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHYE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/154,915
FILING DATE: 19-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,997
FILING DATE: 17-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1993
APPLICATION DATA:
APPLICATION NUMBER: US 07/701,514
FILING DATE: 16-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 13..888
US-08-154-915-1

Query Match 16.5%; Score 216.2; DB 1; Length 1089;
Best Local Similarity 57.1%; Pred. No. 4,1e-61;
Matches 414; Conservative 0; Mismatches 308; Indels 3; Gaps 1;

QY 307 TACTTGAACTTGAGAGCTGGGTGAGGCTTTATGCAAGTTTACAGAGGGGATTAC 366
DB 22 TACGAGAACTGGAAGATTGGGAGGACCTACGAACTGTTCAGGCCAAGAAC 81
QY 367 AGAATTAATGACACATTAAGTCTTAAAGTATCAGCATGAATGCAAGAA--GGA 423
DB 82 CGGAGACATCATGATGCTGTCTTAAACGGGTGAGGCTGATACATGATGAGGT 141
QY 424 GTCCCATTTACAGCTATCGAGAGAGCTTCTCTCGAAGGGTTTGAACATGCAATATT 483
DB 142 GTGCCAGATTCGCCCTCCGGAGATCTGCTTACTCAAGAGGTTGAAGCAAGAACATC 201
QY 484 GTGCTCTCATGATGACATTAATCCACACAAAGACACATGATTCGTTTTGAATCATG 543
DB 202 GTGAGGCTTCATGAGCTCTGACACAGCAAGAAAGCTGATGTTTTTGAATTCGT 261
QY 544 CACACAGACCTGGCCAGATATGTCTCAGCATCCAGAGGGTTTATCTCTAATATGTC 603
DB 262 GACACAGACCTGAGAGATATTTTGAACAGTTGCAATGAGGACCTCATCTCGATTTGA 321
QY 604 AGACTTTTATGTTTCACTTTTGGCGGGCTGGCGATCCACACCAACAGTTCTT 663
DB 322 AAGTATTTCTCTCTTCAGATCTAATAAGGGCTGGATTCTGTATAGCCGATGTGCTA 381
QY 664 CACAGGAGCTGAAACCTGAGAACTTACTATCATGATCACCTGGAGAGCTCAAACTGGCT 723
DB 382 CACAGGAGCTGAAAGCCCGAGAACTGTATTAACAGAAATGGGAGCTGAAATTTGGCT 441
QY 724 GATTTTGGTCTTGCCCGGCGCAAGTCAATTCACAGCAGACATATCTTCAGAAATGCTG 783

DB 442 GATTTTGACCTGCTCGAGCTTTTGGGATTCCTGCTGCTTACTCAGCTAAGGTGTC 501
QY 784 ACCCTCTGTACCGGCCCCCTGATGCTTGTCTGGAGCCAGTAAATTTCTGTAGCTG 843
DB 502 ACATGTGTATCGGCCACCGGATGCTCTTGGGGCCAAAGCTGTATCCAGCTCATC 561
QY 844 GACATATGGGGTGCAGGCTGATCTTTATGAAATGTTCCAGGGTCMACTTTTCT 903
DB 562 GACATGTGTATCGGCCGCTGATCTTTGAGAGCTGGCCAAATGCTGGGGGCTCTTTT 621
QY 904 GGGGTTTCCACATCTTGAACAGCTGAGAAATCTGGAGAGTGTGGAGTCCCTACA 963
DB 622 CCGGCAATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 681
QY 964 GAGATATCTTGGCCGGGAGTCTTCAAGCTACCTAATCAATCAATCAATCAATCAAT 1023
DB 682 GAGAGAGAGTGGCCCTCTATGACCAAGCTGCGCAGACTATTAAGCCCTATCCGATGTA 741
QY 1024 CCTAC 1028
DB 742 GCCAC 746

RESULT 13
US-08-464-517-37
Sequence 37, Application US/08464517
Patent No. 5869640
GENERAL INFORMATION:
APPLICANT: BEACH, David H.
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,517
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1992
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/701,514
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Matthew P. Vincent
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-004C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 13..888

US-08-464-517-37

Query Match 16.5%; Score 216.2; DB 2; Length 1089;
Best Local Similarity 57.1%; Pred. No. 4.1e-61;
Matches 414; Conservative 0; Mismatches 308; Indels 3; Gaps 1;

QY 307 TACTTGAACCTTGAGAGAGCTGGGTGAAAGGCTCTTATGCGACATTTAACAAGGGATTAGC 366
DB 22 TACGAGAACTGGAAGAAAGATTTGGGAGGACCTTACGAACTGTGTTCAAGGCCAAAAAC 81
QY 367 AGAATAAATGACAACAATAAGTGGCTTTAAAGTATCATCATGATGATGAGAGAA---GGA 423
DB 82 CGGAGAGCTCATAGATCGTGGCTCTTAAACGGGTGAGGCTGATGACGATGATGAGGGT 141
QY 424 GTCCCATTTACAGCTATCCGAGAGCTTCTCTCTGAAAGGTTTGAACATGCGCAATTT 483
DB 142 GTGCCGAGTTCGCCCTCCCGGAGATCTGCTCTCAAGAGCTGAAGACACAGAAATC 201
QY 484 GTGCTCCTGATGACATTAATCCACACCAAGAGACATGACATTCGTTTTTGAATACATG 543
DB 202 GTACAGCTTCATACCTCTGACACAGGACAGAGAGCTGACTTTGTTTGAATCTGT 261
QY 544 CACACAGACCTGGCCAGATATATGTCTACAGATCCAGAGGCTTCATCTCATTAATGTC 603
DB 262 GACCAAGAGCTGAAGAATATTTTGAACATGCAATGGTGACCTCGATCTGAGATTGTA 321
QY 604 AGACTTTTCACTTTTCACTTTTGGGGGCTGGCGTACATCCACCAACACAGTTCTT 663
DB 322 AAGTCATTCCTCTTCAGACTTAAAGAGGCTGGGATTCGTCTATGACCGCAATGTGCTA 381
QY 664 CACAGGAGACCTGAAACCTTCAAGACTTACTCATCATGTCACCTGGGAGAGCTCAACCTG 723
DB 382 CACAGGAGACCTGAAAGCCCAAGACCTGCTATTAACAAGAAATGGAGAGCTGAATGGCT 441
QY 724 GATTTTGTCTTCCCGGCGCAAGTCCATTCACGACAGACATCTTTCAGAAATGCTG 783
DB 442 GATTTTGTCTTCCCGCTGAGGCTTTGGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCT 501
QY 784 ACCCTGCTGACCGGCGCTGATGCTTGTGCTGGAGCCCACTGAATATCTCTTGAGCTG 843
DB 502 AACCTGTGTAACCGCCACCGAGTGTCTTGTGGGCGCAAGCTGTAATCCAGTCCATC 561
QY 844 GACATATGAGGAGTCAAGCTGACATCTTATTTGAATGTTCCAGGAGTCAACCTTGTTCCT 903
DB 562 GACATATGAGTCAAGCGGCTGACATCTTGTGAGAGCTGCGCAATGCTGGGCGCTCTTTT 621
QY 904 GGGGTTTCCAACTCTTGAACAGCTGAGAAATGTGGAGGTGCTGGAGTCCCTACA 963
DB 622 CCGGCAATGATGTGATGACCACTTGAAGAGATCTCCGACTGCTGGGAGACGCCACC 681
QY 964 GAGATATCTTGGCCGGAGATCTCCAGCTACCTTAATCAATCCAGAAATGTTCCACTG 1023
DB 682 GAGGAGCACTGGCCCTTATGACCAAGCTGCCAGACTAATAGCCCTATCCGATGTAACCG 741
QY 1024 CCTAC 1028
DB 742 GCCAC 746

RESULT 14
US-08-246-361A-37
Sequence 37, Application US/08246361A
Patent No. 5998582
GENERAL INFORMATION:

APPLICANT: BEACH, David H.
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESSES:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/246,361A
FILING DATE: 19-MAY-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCT-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/888,178
FILING DATE: 26-MAY-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/701,514
FILING DATE: 16-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Mathew P. Vincent
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MIT-004C
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 1089 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURES:
NAME/KEY: CDS
LOCATION: 13..888
US-08-246-361A-37

Query Match 16.5%; Score 216.2; DB 2; Length 1089;
Best Local Similarity 57.1%; Pred. No. 4.1e-61;
Matches 414; Conservative 0; Mismatches 308; Indels 3; Gaps 1;

QY 307 TACTTGAACCTTGAGAGAGCTGGGTGAAAGGCTCTTATGCGACATTTAACAAGGGATTAGC 366
DB 22 TACGAGAACTGGAAGAAAGATTTGGGAGGACCTTACGAACTGTGTTCAAGGCCAAAAAC 81
QY 367 AGAATAAATGACAACAATAAGTGGCTTTAAAGTATCATCATGATGATGAGAGAA---GGA 423
DB 82 CGGAGAGCTCATAGATGTGGCTCTTAAACGGGTGAGGCTGATGACGATGATGAGGGT 141
QY 424 GTCCCATTTACAGCTATCCGAGAGCTTCTCTCTGAAAGGTTTGAACATGCGCAATTT 483
DB 142 GTGCCGAGTTCGCCCTCCCGGAGATCTGCTTCAAGAGCTGAAACACAGAAATCTC 201
QY 484 GTGCTCCTGATGACATTAATCCACCAAGAGACATGACATTCGTTTTTGAATACATG 543
DB 202 GTACAGCTTCATAGAGTCTGCTGACAGGACAGAGAGCTGACTTTGTTTGAATCTGT 261
QY 544 CACACAGACCTGGCCAGATATATGTCTCAGATCCAGAGGCTTCATCTCATTAATGTC 603
DB 262 GACCAAGAGCTGAAGAATATTTTGAACAGTTGCAATGGTGACCTGATCTGAGATTGTA 321
QY 604 AGACTTTTCACTTTTCACTTTTGGGGGCTGGCGTACATCCACCAACACAGTTCTT 663
DB 322 AAGTCATTCCTCTTCAGACTTAAAGAGGCTGGGATTCGTCTATGACCGCAATGTGCTA 381
QY 664 CACAGGAGACCTGAAACCTCAGAACTTACTCATCATGACCTGGAGAGCTCAACCTGAGCT 723
DB 382 CACAGGAGACCTGAAGCCCAAGACCTGCTATTAACAAGAAATGGAGAGCTGAATGGCT 441
QY 724 GATTTTGTCTTCCCGGCGCAAGTCCATTCACGACAGACATCTTTCAGAAATGCTG 783
DB 442 GATTTTGTCTTCCCGCTGAGGCTTTGGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCT 501

QY 784 ACCCTCTGTACCGGCCCCCTGATGCTTGGGAGCCACTGATATATTCCTGAGCTG 843
 DB 502 AACCTGTGTACCGGCCCCGATGCTTGGGAGCCACTGATATATTCCTGAGCTG 843
 QY 844 GACATATGGGAGGAGGCTGATCTTATTTGAATGTCAGAGGTCACCTTTGTTTCT 903
 DB 562 GACATATGGGAGGAGGCTGATCTTATTTGAATGTCAGAGGTCACCTTTGTTTCT 903
 QY 904 GGGGTTTCCACATCTCTGAGAGCTGAGAAATCTGGAGGTCGAGGAGTCCCTACA 963
 DB 622 CCCGGAATGATGTGATGATGACAGATTTGAGAGGATCTTCCAGCTGCTGGGAGCGCCAC 681
 QY 964 GAGGATCTTGGCCGGGAGTCTCCAGCTAATCTAATCAATCAGAAATGTTCCACTG 1023
 DB 682 GAGGAGCACTGGCCCTCTATGACCAAGCTGCCAGACTAATGAGCTTATCCGATGACCG 741
 QY 1024 CCTAC 1028
 DB 742 GTCAC 746

RESULT 15

US-08-463-772-37
 ; Sequence 37, Application US/08463772

Patent No. 6066501

GENERAL INFORMATION:

APPLICANT: BEACH, David H.

TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/463,772

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION NUMBER: US 07/963,308

FILING DATE: 16-OCT-1992

APPLICATION NUMBER: US 07/888,178

FILING DATE: 26-MAY-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/701,514

FILING DATE: 16-MAY-1991

ATTORNEY/AGENT INFORMATION:

NAME: Matthew P. Vincent

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: MI-004C

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 37:

SEQUENCE CHARACTERISTICS:

LENGTH: 1089 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 13..888

US-08-463-772-37

Query Match 16.5%; Score 216.2; DB 3; Length 1089;
 Best Local Similarity 57.1%; Pred. No. 4,1e-61;
 Matches 414; Conservative 0; Mismatches 308; Indels 3; Gaps 1;

QY 307 TACTTAACCTTGAGAGAGCTGGGTGAGGCTCTTATGACACAGTTTACAAAGGATTTAGC 366
 DB 22 TACGAGAACTGAGAAAGATTGGGAGGCACTTACGAACTGTGTCAAGGCCAAAGAAC 81
 QY 367 AGAATTAATGACAACTAGTGGCTTTAAAGTATCAGCATGATGACAGAGAA--GGA 423
 DB 82 CGGAGAGACTCATGAGATCTGGCTTTAAACGGGTGAGGCTGGATGACATGATGAGGCT 141
 QY 424 GTCCATTACAGCTATCCGAGAGCTTCTCTTGAAGGTTTGAATGCCAATATT 483
 DB 142 GTCCGAGTTCCGCTCCGAGAGATCTGCTTACTTAAGAGACTGAGCAAGAACATC 201
 QY 484 GTGCTCTGATGACATTAATCCACCAAGAGACACTGACATTCGTTTTTGAATACATG 543
 DB 202 GTACAGCTTCATGACCTCTGACACAGCAAGAGACTGATCTTTGTTTGAATCTGT 261
 QY 544 CACACAGACTGGCCAGTATATGTCAGCATCCAGAGGAGCTTCATTCATTAATGTC 603
 DB 262 GACAGAGACTGAGAAAGATTTTGAAGTTGCAATGAGTGAACCTCGATCTGAGATTGA 321
 QY 604 AGACTTTTATGTTTCACTTTTGGGGGCTGGGCTATATCCACCAACACGTTCTT 663
 DB 322 AAGTCAATTCCTTCCAGCTACTTAAAGGGCTGGGATTTCTGATGAGCGCAATGTCAT 381
 QY 664 CACAGGAGCTGAGAACTCAGAACTTACTCATCAGTCACTGGGAGAGCTCAAACTGGCT 723
 DB 382 CACAGGAGCTGAGAACTCAGAACTTACTCATCAGTCACTGGGAGAGCTCAAACTGGCT 441
 QY 724 GATTTTGTCTTGGCCGAGGCAAGTCCATTTCCAGCCAGACATCTTTCAGAAAGTCGT 783
 DB 442 GATTTTGGCTGCTGAGCTTTGGGATTTCCGTCGCTGTTACTCAGCTGAGGTGTC 501
 QY 784 ACCCTCTGTACCGGCCCCCTGATGCTTGGGAGCCACTGAAATTCCTGAGCTG 843
 DB 502 AACCTGTGTACCGGCCCCGATGCTTGGGAGCCACTGAAATTCCTGAGCTG 843
 QY 844 GACATATGGGAGGAGGCTGATCTTATTTGAATGTCAGAGGTCACCTTTGTTTCT 903
 DB 562 GACATATGGGAGGAGGCTGATCTTATTTGAATGTCAGAGGTCACCTTTGTTTCT 903
 QY 904 GGGGTTTCCACATCTCTGAGAGCTGAGAAATCTGGAGGTCGAGGAGTCCCTACA 963
 DB 622 CCCGGAATGATGTGATGATGACAGATTTGAGAGGATCTTCCAGCTGCTGGGAGCGCCAC 681
 QY 964 GAGGATCTTGGCCGGGAGTCTCCAGCTAATCTAATCAATCAGAAATGTTCCACTG 1023
 DB 682 GAGGAGCACTGGCCCTCTATGACCAAGCTGCCAGACTAATGAGCTTATCCGATGACCG 741
 QY 1024 CCTAC 1028
 DB 742 GTCAC 746

Search completed: December 27, 2004, 11:08:27
 Job time : 70.1305 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2004, 17:40:11 ; Search time 2501.82 Seconds
(without alignments) 19051.434 Million cell updates/sec

Title: US-10-786-065-4

Perfect score: 1308

Sequence: 1 atgggtcaagcgtcgtgc.....agttacgaacgtcgtgta 1308

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	687	52.5	804	4	B1755983	B1755983 603030316
2	652.4	49.9	766	2	BE562611	BE562611 601336446
3	592.8	45.3	602	6	CD636753	CD636753 56015009H
4	591.6	45.2	598	6	CD636752	CD636752 56015001J
5	583.2	44.6	601	6	CD636760	CD636760 56015089J
6	580.8	44.4	600	6	CD636754	CD636754 56015009J
7	578.6	44.2	602	6	CD636748	CD636748 56014973J
8	574.2	43.9	602	6	CD636751	CD636751 56015001H
9	573.6	43.9	596	6	CD636758	CD636758 56015025J
10	570.4	43.6	598	6	CD636740	CD636740 56014901J
11	566.6	43.3	602	6	CD636757	CD636757 56015025H
12	565.8	43.3	602	6	CD636759	CD636759 56015089H
13	564.2	43.1	601	6	CD636746	CD636746 56014925J
14	560	42.8	602	6	CD636755	CD636755 56015017H
15	557.2	42.6	587	6	CD636750	CD636750 56014981J
16	557.2	42.6	598	6	CD636756	CD636756 56015017J
17	536.4	41.0	586	6	CD636749	CD636749 56014981H
18	534.6	40.9	603	6	CD636739	CD636739 56014901H
19	520.6	39.8	603	6	CD636742	CD636742 56014903J
20	514.8	39.4	600	6	CD636741	CD636741 56014909H
21	512.4	39.2	538	6	CD636744	CD636744 56014917J
22	512.4	39.2	599	6	CD636747	CD636747 56014973H
23	500.2	38.2	623	4	BM539173	BM539173 hb05f12.9
24	481	36.8	481	1	AA436054	AA436054 zt01c12.1

25	476	36.4	499	5	BX283916	BX283916
26	472.4	36.1	586	6	CD636745	CD636745 56014925H
27	444.6	34.0	660	4	BD326162	BD326162 602425156
28	388.2	29.7	540	9	CG523756	CG523756 OST96514
29	370.2	28.3	3067	3	AK087398	AK087398 Mus muscu
30	367	28.1	2435	3	AK083269	AK083269 Mus muscu
31	348.2	26.6	745	6	BI345276	BI345276 UI-M-GHO-
32	346.2	26.5	580	4	BI452376	BI452376 374050 MA
33	340.8	26.1	471	4	BO720115	BO720115 602691042
34	333.4	25.5	471	4	AK051283	AK051283 Mus muscu
35	332	25.4	468	4	BG772738	BG772738 602720896
36	327.4	25.0	729	6	CB525597	CB525597 UI-M-FYO-
37	325.4	24.9	1322	9	AY406869	AY406869 Homo sap1
38	325.4	24.9	879	5	BO220223	BO220223 AGENCOURT
39	323.6	24.7	731	6	CF536868	CF536868 UI-M-FYO-
40	319.4	24.7	762	6	CB523552	CB523552 UI-M-GHO-
41	318.2	24.3	1290	9	AY406870	AY406870 Pan trogl
42	318	24.3	633	6	CB522451	CB522451 UI-M-GHO-
43	316.8	24.2	598	5	BU265800	BU265800 603814271
44	316.4	24.2	786	7	CN537781	CN537781 UI-M-HSO-
45	314.8	24.1	929	5	BU464167	BU464167 603366128

ALIGNMENTS

RESULT 1
LOCUS B1755983 804 bp mRNA linear EST 25-SEP-2001
DEFINITION 603030316F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200755 5',
mRNA sequence.
ACCESSION B1755983
VERSION B1755983.1 GI:15747561
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at:
http://image.llnl.gov
Plate: LLM11502 row: p column: 04
High quality sequence stop: 797.
Location/Qualifiers

FEATURES
source 1..804
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5200755"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_114"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains; Age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 52.5%; Score 687; DB 4; Length 804;

Best Local Similarity 97.5%; Pred. No. 2.5e-196;
Matches 740; Conservative 0; Mismatches 15; Indels 4; Gaps 4;

QY 1 ATGGGTCAAGAGCTGTGTGCAAAAGCTGTATACACCTGTGATGACGTGTCTACATTTGTCA 60
DB 45 ATGGGTCAAGAGCTGTGTGCAAAAGCTGTATACACCTGTGATGACGTGTCTACATTTGTCA 104
QY 61 GAGGAGGCGGAGGCAACACAGCTGTGAGAGAGTCAAGCTGTGACACACGAGGCTGTGCTTC 120
DB 105 GAGGAGGCGGAGGCAACACAGCTGTGAGAGAGTCAAGCTGTGACACACGAGGCTGTGCTTC 164
QY 121 AAGCTTAACAGACCTTAAGAGAGCATGTGTTCATGATCTTCAATTCACCCCAAGGAGCTT 180
DB 165 AAGCTTAACAGACCTTAAGAGAGCATGTGTTCATGATCTTCAATTCACCCCAAGGAGCTT 224
QY 181 CAAGCTGCCGTGCGCAGAAATTCAAGATTAAGAGCCACGAGTAACAGTATTTT 240
DB 225 CAAGCTGCCGTGCGCAGAAATTCAAGATTAAGAGCCACGAGTAACAGTATTTT 284
QY 241 CAGGAAGAGATTTGAGGCGAGGTTTTCACTGAGAGAGAGCTCCCTTTTGGGCGACCC 300
DB 285 CAGGAAGAGATTTGAGGCGAGGTTTTCACTGAGAGAGAGCTCCCTTTTGGGCGACCC 344
QY 301 TCATCTTACTGTAACTTGAGAGAGCTGGGTGAAGGCTTTATGSCACAGTTTACAAGGAG 360
DB 345 TCATCTTACTGTAACTTGAGAGAGCTGGGTGAAGGCTTTATGSCACAGTTTACAAGGAG 404
QY 361 ATTGAGAGAAATATGAGCAACTAGTGGCTTTAA-AGTCATCAGCATGATGCAAGAGA 419
DB 405 ATTGAGAGAAATATGAGCAACTAGTGGCTTTAA-AGTCATCAGCATGATGATGCAAGAGA 464
QY 420 AAGAGTCCATTTTAAAGCTATCCGAGAGCTTCTCTGAGAGGTTTGAACATGCCAA 479
DB 465 AAGAGTCCATTTTAAAGCTATCCGAGAGCTTCTCTGAGAGGTTTGAACATGCCAA 524
QY 480 TATTGTGCTCTGTCATGA-CATATCACACCAAGAGACATGACATTGGTTTTGAAT 538
DB 525 TATTGTGCTCTGTCATGA-CATATCACACCAAGAGACATGACATTGGTTTTGAAT 584
QY 539 AC-ATGCACACAGACCTGGGCCAGATATATGTCTCAGCATTCAGAGAGGCTTCAATCCAT 597
DB 585 ACATGCACACAGACCTGGGCCAGATATATGTCTCAGCATTCAGAGAGGCTTCAATCCAT 644
QY 598 AATGTGAGCTTTTCAATGTTTCACTTTTGGGGGCTTGCGTATCCACCAACACAC 657
DB 645 AATGTGAGCTTTTCAATGTTTCACTTTTGGGGGCTTGCGTATCCACCAACACAC 704
QY 658 GTTCTTCAAGGAGACCTGAAACCTCAGAACTTATCATGATCACTT-GGGAGAGCTCAA 716
DB 705 GTTCTTCAAGGAGACCTGAAACCTCAGAACTTATCATGATCACTTGGGAGAGCTCAA 764
QY 717 ACTGCTGATTTTGTGCTTGGCCGGGCGCAAGTCAATCC 755
DB 765 ACTGCTGATTTTGTGCTTGGCCGGGCGCAAGTCAATCC 803

RESULT 2
BE562611 766 bp mRNA linear EST 15-AUG-2000
LOCUS 60136446F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3690395 5',
DEFINITION mRNA sequence.
ACCESSION BE562611
VERSION BE562611.1 GI:9806331
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 766)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://Image.lnl.gov
Plate: LNC386 row: 12
High quality sequence stop: 714.
Location/Qualifiers
1. 766
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3690395"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_44"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
SmaI; cDNA made by Oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
Query Match 49.9%; Score 652.4; DB 2; Length 766;
Best Local Similarity 97.6%; Pred. No. 7.7e-186;
Matches 736; Conservative 0; Mismatches 11; Indels 7; Gaps 7;

QY 123 GCTAACAGACTTAAGAAAGCATCATGTTCCATGATCTTCAATTCACCCAGGGGACTTCA 182
DB 19 GCTAACAGACTTAAGAAAGCATCATGTTCCATGATCTTCAATTCACCCAGGGGACTTCA 77
QY 183 AGCTGCCGTGCGCAGAAATTCAAGATTAAGAGCCAGGAGTAACAGTATTTTCA 242
DB 78 AGCTGCCGTGCGCAGAAATTCAAGATTAAGAGCCAGGAGTAACAGTATTTTCA 136
QY 243 GGAAGAGATCTGAGGCGAGGTTTTCACTGAGAGAGAGCTCCCTTTTGGGCGACCTC 302
DB 137 GGAAGAGATCTGAGGCGAGGTTTTCACTGAGAGAGAGCTCCCTTTTGGGCGACCTC 195
QY 303 ATCTTACTTGAACCTTGAGAGAGCTGGTGAAGCTCTTATGCGACATTTACAGGGGAT 362
DB 196 ATCTTACTTGAACCTTGAGAGAGCTGGTGAAGCTCTTATGCGACATTTACAGGGGAT 255
QY 363 TAGCAGATTAATGAGCAACTAGTGGCTTTAAAGTCATCAGATGAGAGAGAG 422
DB 256 TAGCAGATTAATGAGCAACTAGTGGCTTTAAAGTCATCAGATGAGAGAGAGAG 315
QY 423 AGTCCATTTTACAGCTATCCGAGAGCTTCTCTCGTGAAGGCTTGAACATGCCAATAT 482
DB 316 AGTCCATTTTACAGCTATCCGAGAGCTTCTCTCGTGAAGGCTTGAACATGCCAATAT 375
QY 483 TGTGCTCTGTCATGATATATTCACACCAAGAGACCTGATTCGTTTGAATATAT 542
DB 376 TGTGCTCTGTCATGATATATTCACACCAAGAGACCTGATTCGTTTGAATATAT 434
QY 543 GCACACAGACCTGGGCCAGATATATGTCTCAGATCCAGAGGGCTTCAATCCATTAATGT 602
DB 435 GCACACAGACCTGGGCCAGATATATGTCTCAGATCCAGAGGGCTTCAATCCATTAATGT 494
QY 603 CAGACTTTTCACTGTTTCACTTTTGGGGGCTTGCGTATCCACCAACAGCTTCT 662
DB 495 CAGACTTTTCACTGTTTCACTTTTGGGGGCTTGCGTATCCACCAACAGCTTCT 554
QY 663 TCACAGGAGACTGAAGACTCAGAACTTATCTATCATGATCACTGGGAGAGCTCAAGCTGCC 722
DB 555 TCACAGGAGACTGAAGACTCAGAACTTATCTATCATGATCACTGGGAGAGCTCAAGCTGCC 614
QY 723 TGATTTTGTGCTTGGCCGGGCGCAAGTCAATTCACGACGACATCTTTCAGAAAGTGAT 782

```

Db      615 TGAATTTGATCTTGGCCGGGACC-AGTCCATTTCCACCGACGACATCTTTCAGAAATGCT 673
Qy      783 GACCCCTGTGTACCGGCCCCCTGTATGCTTGTCTGGAGCCATGAAATATTCCTGTAGCT 842
Db      674 GAACCTCTGTGTA-CGGCCCCCTGTATGC-TTGTCTGGAGCCATGAAATATTCCTGTAGCT 731
Qy      843 GGACATATGGGGTGCAGGCTGCATCTTTATTGAA 876
Db      732 GGACCTATGGGGTGCACGCTGTCTTATTATGCAA 765

RESULT 3
CD636753      602 bp      mRNA      linear      EST 12-JAN-2004
LOCUS         56015009H1 FLP Homo sapiens cDNA, mRNA sequence.
DEFINITION   CD636753
ACCESSION    CD636753.1 GI:40285020
VERSION      EST.
KEYWORDS     Homo sapiens (human)
SOURCE       Homo sapiens
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 602)
AUTHORS      Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
TITLE        Circular rapid amplification of cDNA ends for high-throughput
              Genomics 84 (1), 205-210 (2004)
JOURNAL      Contact: Fu GK
COMMENT      Incyte Genomics, Inc.
              3160 Porter Dr., Palo Alto, CA 94304, USA
              Tel: 6508454102
              Email: gfu@incyte.com.
              Location/Qualifiers
FEATURES     source
              1..602
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone_1fb="FLP"
               /note="Vector: pDrive Cloning Vector"
ORIGIN
Query Match      45.3%; Score 592.8; DB 6; Length 602;
Best Local Similarity 99.7%; Pred. No. 8.3e-168;
Matches 594; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

410 ATGCAGAGGAGAGATCCCATTTACAGTATCCGAGAGCTTCTCTCGTGAAGGTTTCA 469
Db      6 AGGCGAGAGAGAGATCCCATTTACAGTATCCGAGAGCTTCTCTCGTGAAGGTTTCA 65
Qy      470 AACATGCCAATATTTGTCTCTGCATGACATATTCACACCAAGAGACATGACATTCG 529
Db      66 AACATGCCAATATTTGTCTCTGCATGACATATTCACACCAAGAGACATGACATTCG 125
Qy      530 TTTTGTGATACATGACACAGACCTGGCCAGTATATGTCTCAGCATCCAGAGGGCTTC 589
Db      126 TTTTGAATACATGACACAGACCTGGCCAGTATATGTCTCAGCATCCAGAGGGCTTC 185
Qy      590 ATTCCTCATATATGTCAGACTTTTCATGTTTCACTTTTGGGGGGCCGTGGGTATCATCCACC 649
Db      186 ATTCCTCATATATGTCAGACTTTTCATGTTTCACTTTTGGGGGGCCGTGGGTATCATCCACC 245
Qy      650 ACCAACACGTTTCTTACAGAGGACCTGAAACCTCAGAACTTACTCATCACTGAGAG 709
Db      246 ACCAACACGTTTCTTACAGAGGACCTGAAACCTCAGAACTTACTCATCACTGAGAG 305
Qy      710 AGCTCAAAATGCTGATTTTGTGCTTGGCCGGGCGCAAGTCCATTCGACGCACTACT 769
Db      306 AGCTCAAAATGCTGATTTTGTGCTTGGCCGGGCGCAAGTCCATTCGACGCACTACT 365
Qy      770 CTTCAGAAAGTCTGACCTCTGTGTACCGGCCCCCTGTATGCTTGTCTGGAGCCACTGAAT 829
Db      366 CTTCAGAAAGTCTGACCTCTGTGTACCGGCCCCCTGTATGCTTGTCTGGAGCCACTGAAT 425

```

```

Qy      830 ATTCTCTGAGCTGACATATGCGGTGCAGGCTGCATCTTTATTTGAATATGTTCCAGGCTC 889
Db      426 ATTCTCTGAGCTGACATATGCGGTGCAGGCTGCATCTTTATTTGAATATGTTCCAGGCTC 485
Qy      890 AACCTTTGTTCTCGGGGTTTCCACATCTTGAAACAGCTGAGAAATCTGGAGAGTGC 949
Db      486 AACCTTTGTTCTCGGGGTTTCCACATCTTGAAACAGCTGAGAAATCTGGAGAGTGC 545
Qy      950 TGGAGTCCCTACAGAGATACCTTGCCCGGAGTCTCCAACTACTTAACAAT 1005
Db      546 TGGAGTCCCTACAGAGATACCTTGCCCGGAGTCTCCAACTACTTAACAAT 601

RESULT 4
CD636752      598 bp      mRNA      linear      EST 12-JAN-2004
LOCUS         56015001J1 FLP Homo sapiens cDNA, mRNA sequence.
DEFINITION   CD636752
ACCESSION    CD636752
VERSION      CD636752.1 GI:40285019
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 598)
AUTHORS      Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
TITLE        Circular rapid amplification of cDNA ends for high-throughput
              Genomics 84 (1), 205-210 (2004)
JOURNAL      Contact: Fu GK
COMMENT      Incyte Genomics, Inc.
              3160 Porter Dr., Palo Alto, CA 94304, USA
              Tel: 6508454102
              Email: gfu@incyte.com.
              Location/Qualifiers
FEATURES     source
              1..598
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone_1fb="FLP"
               /note="Vector: pDrive Cloning Vector"
ORIGIN
Query Match      45.2%; Score 591.6; DB 6; Length 598;
Best Local Similarity 99.3%; Pred. No. 1.9e-167;
Matches 594; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

408 GAATGACAGAGAGAGATCCCATTTACAGTATCCGAGAGCTTCTCTCTGAAAGGTTT 467
Db      598 GAGCGCAGAGAGAGAGATCCCATTTACAGTATCCGAGAGCTTCTCTCTGAAAGGTTT 539
Qy      468 GAAATGCCAATATTTGTCTCTGCATGACATATTCACACCAAGAGACATGACATTC 527
Db      538 GAAATGCCAATATTTGTCTCTGCATGACATATTCACACCAAGAGACATGACATTC 479
Qy      538 CGTTTGTGAATACATGACACAGACCTGGCCAGTATATGTCTCAGCATCCAGAGGGCT 587
Db      478 CGTTTGTGAATACATGACACAGACCTGGCCAGTATATGTCTCAGCATCCAGAGGGCT 419
Qy      588 TATCTCTCATATATGTCAGACTTTTCATGTTTCACTTTTGGGGGGCTGGGCTATATCCA 647
Db      418 TATCTCTCATATATGTCAGACTTTTCATGTTTCACTTTTGGGGGGCTGGGCTATATCCA 359
Qy      648 CCACCAACAGTTTCTTACAGAGGACCTGAAACCTCAGAACTTACTCATCACTGACCTGG 707
Db      358 CCACCAACAGTTTCTTACAGAGGACCTGAAACCTCAGAACTTACTCATCACTGACCTGG 299
Qy      708 AGAGCTCAAACTGAGTGAATTTGTGCTTGGCCGGGCGCAAGTCCATTCCTCCAGCAGACATA 767
Db      298 AGAGCTCAAACTGAGTGAATTTGTGCTTGGCCGGGCGCAAGTCCATTCCTCCAGCAGACATA 239
Qy      768 CTCTTCAGAAAGTCTGACCTCTGTGTACCGGCCCCCTGTATGCTTGTCTGGAGCCACTGA 827

```

```
Db 238 CTCCTTGAAGAGTCGTAACCCCTCTGTAACGGCCCCCTGATGCTTCTGGAGCCCACTGA 179
|||
Qy 828 ATATTCCTCTGAGCTGAGCATATGAGGCTGAGCTGCATCTTATTAATGATGTTCCAGG 887
Db 178 ATATTCCTCTGAGCTGAGCATATGAGGCTGAGCTGCATCTTATTAATGATGTTCCAGG 119
Qy 888 TCACCTCTGTTGTTCTCTGGGGTTTCCAAATCCTTGAACAGCTGAGAAATCTGGAGGT 947
Db 118 TCACCTCTGTTGTTCTCTGGGGTTTCCAAATCCTTGAACAGCTGAGAAATCTGGAGGT 59
Qy 948 GCTGGAGTCCCTACAGAGATCTTGGCCGGGAGTCTCCAGCTACTACTACTAAT 1005
Db 58 GCTGGAGTCCCTACAGAGATCTTGGCCGGGAGTCTCCAGCTACTACTACTAAT 1

RESULT 5
CD636760 601 bp mRNA linear EST 12-JAN-2004
LOCUS CD636760
DEFINITION CD636760 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD636760.1 GI:40285027
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
AUTHORS 1 (bases 1 to 601)
TITLES Circular rapid amplification of cDNA ends for high-throughput
JOURNAL extension cloning of partial genes
COMMENT Genomics 84 (1), 205-210 (2004)
CONTACT: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
FEATURES
source
1..601
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN
Query Match 44.6%; Score 583.2; DB 6; Length 601;
Best Local Similarity 99.5%; Pred. No. 6.8e-165;
Matches 585; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 410 ATGCAGAGGAGAGGATCCCATTTACAGCTATCCGAGAGCTTCTCTGAGAGGTTTGA 469
Db 6 AGGCAAGAGGAGAGGATCCCATTTACAGCTATCCGAGAGCTTCTCTGAGAGGTTTGA 65
Qy 470 AACATGCCAATATTTGCTCTGCAATGATATTCACACCAAGAGACATGACATTCG 529
Db 66 AACATGCCAATATTTGCTCTGCAATGATATTCACACCAAGAGACATGACATTCG 125
Qy 530 TTTTGAATACATGACAGACAGCTGCGCCAGTATATGCTCAGCATCCAGAGAGGCTTC 589
Db 126 TTTTGAATACATGACAGACAGCTGCGCCAGTATATGCTCAGCATCCAGAGAGGCTTC 185
Qy 590 ATCTCATATATGTCAGACATTTTCATGTTTCACTTTTGGGGGCTGGGGTATCATCCACC 649
Db 186 ATCTCATATATGTCAGACATTTTCATGTTTCACTTTTGGGGGCTGGGGTATCATCCACC 245
Qy 650 ACCAACAAGTTCTTCAAGGAGCTGAAACCTCAGAACTTACTCATCAGTCACTGGAG 709
Db 246 ACCAACAAGTTCTTCAAGGAGCTGAAACCTCAGAACTTACTCATCAGTCACTGGAG 305
Qy 710 AGCTCAACTGGCTGATTTTGGTCTTGCCCGGGCCAAAGTCAATCCACGACAGACATACT 769
Db 306 AGCTCAACTGGCTGATTTTGGTCTTGCCCGGGCCAAAGTCAATCCACGACAGACATACT 365
```

```
Qy 770 CTTCAAGAGTGTGACACCTCTGATACCGCCCCCTGATGCTTGTGAGGAGCCCACTGAT 829
Db 366 CTTCAAGAGTGTGACACCTCTGATACCGCCCCCTGATGCTTGTGAGGAGCCCACTGAT 425
Qy 830 ATTCCCTGAGCTGAGCATATGAGGAGGAGCTGCATCTTATTAATGATGTTCCAGGCTC 889
Db 426 ATTCCCTGAGCTGAGCATATGAGGAGGAGCTGCATCTTATTAATGATGTTCCAGGCTC 485
Qy 890 AACCTTTGTTTCTGCGGGTTTCCAAATCTCTTGAACAGCTGAGAAATCTGGAGAGTGC 949
Db 486 AACCTTTGTTTCTGCGGGTTTCCAAATCTCTTGAACAGCTGAGAAATCTGGAGAGTGC 545
Qy 950 TGGAGTCCCTACAGAGATACTTGGCCGGGAGTCTCCAGCTACTCTA 997
Db 546 TGGAGTCCCTACAGAGATACTTGGCCGGGAGTCTCCAGCTACTCTA 593

RESULT 6
CD636754/c 600 bp mRNA linear EST 12-JAN-2004
LOCUS CD636754/c
DEFINITION CD636754 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD636754
VERSION CD636754.1 GI:40285021
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
AUTHORS 1 (bases 1 to 600)
TITLES Circular rapid amplification of cDNA ends for high-throughput
JOURNAL extension cloning of partial genes
COMMENT Genomics 84 (1), 205-210 (2004)
CONTACT: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
FEATURES
source
1..600
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN
Query Match 44.4%; Score 580.8; DB 6; Length 600;
Best Local Similarity 99.5%; Pred. No. 3.6e-164;
Matches 593; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 410 ATGCAGAGGAGAGATCCCATTTACAGCTATCCGAGAGCTTCTCTGAGAGGTTTGA 469
Db 595 AGGCAAGAGGAGAGATCCCATTTACAGCTATCCGAGAGCTTCTCTGAGAGGTTTGA 536
Qy 470 AACATGCCAATATTTGCTCTGCAATGATATTCACACCAAGAGACATGACATTCG 529
Db 535 AACATGCCAATATTTGCTCTGCAATGATATTCACACCAAGAGACATGACATTCG 476
Qy 530 TTTTGAATACATGACAGACAGCTGCGCCAGTATATGCTCAGATCCAGAGAGGCTTC 589
Db 475 TTTTGAATACATGACAGACAGCTGCGCCAGTATATGCTCAGATCCAGAGAGGCTTC 416
Qy 590 ATCTCATATATGTCAGACATTTTCATGTTTCACTTTTGGGGGCTGGGGTATCATCCACC 649
Db 415 ATCTCATATATGTCAGACATTTTCATGTTTCACTTTTGGGGGCTGGGGTATCATCCACC 356
Qy 650 ACCAACAAGTTCTTCAAGGAGCTGAAACCTCAGAACTTACTCATCAGTCACTGGAG 709
Db 355 ACCAACAAGTTCTTCAAGGAGCTGAAACCTCAGAACTTACTCATCAGTCACTGGAG 296
Qy 710 AGCTCAACTGGCTGATTTTGGTCTTGCCCGGGCCAAAGTCAATCCACGACAGACATACT 769
```

Db 295 AGCTCAAGTGGTGAATTTTGGTCTTGGCCCGGCAAGTCCATTTCCAGCCAGATCT 236
Qy 770 CTTCAGAAAGTGTGACCTCTGTGTACCGGCCCCCTGATCTTGTCTGGAGCCACTGAAT 829
Db 235 CTTCAGAAAGTGTGACCTCTGTGTACCGGCCCCCTGATCTTGTCTGGAGCCACTGAAT 176
Qy 830 ATTCTCTGAGCTGACATATGGGGTGCAGGGCTGCATCTTTATTTGAATGTTCCAGGGTTC 889
Db 175 ATTCTCTGAGCTGACATATGGGGTGCAGGGCTGCATCTTTATTTGAATGTTCCAGGGTTC 116
Qy 890 AACCTTTGTTCTCTGGGGTTTCCACATCTTGAAGCTGGAAGAAATCTGGAGAGTGC 949
Db 115 AACCTTTGTTCTCTGGGGTTTCCACATCTTGAAGCTGGAAGAAATCTGGAGAGTGC 56
Qy 950 TGGGAGTCCCTTACAGAGATACCTTGGCCGGAGTCTCCAAGCTACTTAACAAT 1005
Db 55 TGGGAGTCCCTTACAGAGATACCTTGGCCGGAGTCTCCAAGCTACTTAACAAT 1

RESULT 7
CD636748/c 602 bp mRNA linear EST 12-JAN-2004
LOCUS CD636748 56014973J1 FLP Homo sapiens cDNA, mRNA sequence.
DEFINITION CD636748.1 GI:40285015
ACCESSION CD636748.1 GI:40285015
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE EST.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 602)
AUTHORS Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.

FEATURES
source Location/Qualifiers
1..602
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN
Query Match 44.2%; Score 578.6; DB 6; Length 602;
Best Local Similarity 99.2%; Pred. No. 1.7e-163;
Matches 592; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
Qy 410 ATGCAGAGAGAGAGTCCCATTTACGCTATCCGAGAGCTTCTCTCGAAGGGTTTGA 469
Db 597 AGCGAGAGAGAGAGTCCCATTTACGCTATCCGAGAGCTTCTCTCGAAGGGTTTGA 538
Qy 470 AACATGCCAATATTTGTGCTCTGCAATGAGATATCCAGACC-AAAGAGACATGACATTC 528
Db 537 AACATGCCAATATTTGTGCTCTGCAATGAGATATCCAGACC-AAAGAGACATGACATTC 478
Qy 539 GTTTTGAATATGATGACACAGACCTGGCCAGATATATGCTCAGATCCAGAGGGCTT 588
Db 477 GTTTTGAATATGATGACACAGACCTGGCCAGATATATGCTCAGATCCAGAGGGCTT 418
Qy 589 CATCTCATATATGTCAGACTTTTCATGTTTCACTTTTGGGGGCTTGGCGTATCCAC 648
Db 417 CATCTCATATATGTCAGACTTTTCATGTTTCACTTTTGGGGGCTTGGCGTATCCAC 358
Qy 649 CACCAACAGCTTCTTACAGAGGACCTGAACCTCAGAACTTACTATCAGTACACCTGGGA 708
Db 357 CACCAACAGCTTCTTACAGAGGACCTGAACCTCAGAACTTACTATCAGTACACCTGGGA 298

Qy 709 GAGCTCAAACTGGCTGATTTTGGTCTTGGCCCGGCAAGTCCATTTCCAGCCAGATAC 768
Db 297 GAGCTCAAACTGGCTGATTTTGGTCTTGGCCCGGCAAGTCCATTTCCAGCCAGATAC 238
Qy 769 TCTTCAGAAAGTGTGACCTCTGTGTACCGGCCCCCTGATCTTGTCTGGAGCCACTGA 828
Db 237 TCTTCAGAAAGTGTGACCTCTGTGTACCGGCCCCCTGATCTTGTCTGGAGCCACTGA 178
Qy 829 TATCTCTGAGCTGACATATGGGGTGCAGGGCTGCATCTTTATTTGAATGTTCCAGGGT 888
Db 177 TATCTCTGAGCTGACATATGGGGTGCAGGGCTGCATCTTTATTTGAATGTTCCAGGGT 118
Qy 889 CAACCTTTGTTCTCTGGGGTTTCCACATCTTGAACAGCTGGAAGAAATCTGGAGAGTGC 948
Db 117 CAACCTTTGTTCTCTGGGGTTTCCACATCTTGAACAGCTGGAAGAAATCTGGAGAGTGC 58
Qy 949 CTGGGAGTCCCTTACAGAGATACCTTGGCCGGAGTCTCCAAGCTACTTAACAAT 1005
Db 57 CTGGGAGTCCCTTACAGAGATACCTTGGCCGGAGTCTCCAAGCTACTTAACAAT 1

RESULT 8
CD636751 602 bp mRNA linear EST 12-JAN-2004
LOCUS CD636751 56015001H1 FLP Homo sapiens cDNA, mRNA sequence.
DEFINITION CD636751
ACCESSION CD636751.1 GI:40285018
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE EST.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 602)
AUTHORS Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.

FEATURES
source Location/Qualifiers
1..602
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN
Query Match 43.9%; Score 574.2; DB 6; Length 602;
Best Local Similarity 99.2%; Pred. No. 3.6e-162;
Matches 587; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
Qy 410 ATGCAGAGAGAGTCCCATTTACGCTATCCGAGAGCTTCTCTCGAAGGGTTTGA 469
Db 6 AGCGAGAGAGAGTCCCATTTACGCTATCCGAGAGCTTCTCTCGAAGGGTTTGA 65
Qy 470 AACATGCCAATATTTGTGCTCTGCAATGAGATATCCAGACC-AAAGAGACATGACATTC 529
Db 66 AACATGCCAATATTTGTGCTCTGCAATGAGATATCCAGACC-AAAGAGACATGACATTC 478
Qy 539 GTTTTGAATATGATGACACAGACCTGGCCAGATATATGCTCAGATCCAGAGGGCTT 589
Db 126 GTTTTGAATATGATGACACAGACCTGGCCAGATATATGCTCAGATCCAGAGGGCTT 185
Qy 590 ATCTCATATATGTCAGACTTTTCATGTTTCACTTTTGGGGGCTTGGCGTATCCAC 649
Db 186 ATCTCATATATGTCAGACTTTTCATGTTTCACTTTTGGGGGCTTGGCGTATCCAC 245
Qy 650 ACCAACAAGCTTCTTACAGAGGACCTGAACCTCAGAACTTACTATCAGTACACCTGGGA 709

Db 246 ACCAACAAGCTTCTACAGAGGACCTGAACCTCAGAACTTACCCATCAGACCTGGAG 305
Qy 710 AGCTCAACTGCGTGAATTTTGTCTTGGCCCGGCAAGTCATTTCCAGCAGACATACT 769
Db 306 AGCTCAACTGCGTGAATTTTGTCTTGGCCCGGCAAGTCATTTCCAGCAGACATACT 365
Qy 770 CTTCAGAAAGTCGTCAGACCTTGTGTACCGGCCCCCTGATGCTTGTGGAGCACTGAAT 829
Db 366 CTTCAGAAAGTCGTCAGACCTTGTGTACCGGCCCCCTGATGCTTGTGGAGCACTGAAT 425
Qy 830 ATTCTCTGAGCTGACATATGAGGAGTCGAGGCTGACATCTTATTTGAATTTCAAGGTC 889
Db 426 ATTCTCTGAGCTGACATATGAGGAGTCGAGGCTGACATCTTATTTGAATTTCAAGGTC 485
Qy 890 AACCTTTGTTCTCTGGGGTTTCCACATCTTGAACAGCTGAGAAAATCTGGAGAGTC 949
Db 486 AACCTTTGTTCTCTGGGGTTTCCACATCTTGAACAGCTGAGAAAATCTGGAGAGTC 545
Qy 950 TGGAGTCCCTACAGAGATCTTGGCCGGAGTCTTCAAGCTTAACTA 1001
Db 546 TGGAGTCCCTACAGAGATCTTGG-CCGGGGTCTTCAAGCTTAACTA 596

RESULT 9
CD636758/c 596 bp mRNA linear EST 12-JAN-2004
LOCUS CD636758
DEFINITION 56015025J1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD636758
VERSION CD636758.1 GI:40285025
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 596)
AUTHORS Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
JOURNAL extension cloning of partial genes
COMMENT Genomics 84 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers
FEATURES
source 1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_idb="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN
Query Match 43.9%; Score 573.6; DB 6; Length 596;
Best Local Similarity 99.2%; Pred. No. 5.5e-162;
Matches 587; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 650 ACCAACAAGCTTCTACAGAGGACCTGAACCTCAGAACTTACTCATCACTCAGGAG 709
Db 351 ACCAACAAGCTTCTACAGAGGACCTGAACCTCAGAACTTACTCATCACTCAGGAG 292
Qy 710 AGCTCAACTGCGTGAATTTTGTCTTGGCCCGGCAAGTCATTTCCAGCAGACATACT 769
Db 291 AGCTCAACTGCGTGAATTTTGTCTTGGCCCGGCAAGTCATTTCCAGCAGACATACT 232
Qy 770 CTTCAGAAAGTCGTCAGACCTTGTGTACCGGCCCCCTGATGCTTGTGGAGCACTGAAT 829
Db 231 CTTCAGAAAGTCGTCAGACCTTGTGTACCGGCCCCCTGATGCTTGTGGAGCACTGAAT 172
Qy 830 ATTCTCTGAGCTGACATATGAGGAGTCGAGGCTGACATCTTATTTGAATTTTCAAGGTC 889
Db 171 ATTCTCTGAGCTGACATATGAGGAGTCGAGGCTGACATCTTATTTGAATTTTCAAGGTC 112
Qy 890 AACCTTTGTTCTCTGGGGTTTCCACATCTTGAACAGCTGAGAAAATCTGGAGAGTC 949
Db 111 AACCTTTGTTCTCTGGGGTTTCCACATCTTGAACAGCTGAGAAAATCTGGAGAGTC 52
Qy 950 TGGAGTCCCTACAGAGATCTTGGCCGGAGTCTTCAAGCTTAACTA 1001
Db 51 TGGAGTCCCTACAGAGATCTT-GCCAGAGTCTTCAAACTAACTA 51

RESULT 10
CD636740/c 598 bp mRNA linear EST 12-JAN-2004
LOCUS CD636740
DEFINITION 56014901J1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD636740
VERSION CD636740.1 GI:40285007
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 598)
AUTHORS Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
JOURNAL extension cloning of partial genes
COMMENT Genomics 84 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers
FEATURES
source 1..598
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_idb="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN
Query Match 43.6%; Score 570.4; DB 6; Length 598;
Best Local Similarity 98.5%; Pred. No. 5.2e-161;
Matches 587; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

```

Db      413 ATCTCATATATGTCAGACCTTTTCATGTTTCACTTTTGCGGGGCTGGGCTACATCCACC 354
Qy      650 ACCAACACGTTTCTTCAAGAGGACCTGAAAACCTGAACTTACTCATCAGTCACTGGAG 709
Db      353 ACCAACACGTTTCTTCAAGAGGACCTGAAAACCTGAACTTACTCATCAGTCACTGGAG 294
Qy      710 AGCTCAAACTGGCTGATTTTGTCTTGGTCCCGGGCCAAAGTCCATTTCCAGCAACATTA 769
Db      293 AGCTCAAACTGGCTGATTTTGTCTTGGTCCCGGGCCAAAGTCCATTTCCAGCAACATTA 234
Qy      770 CTTCAGAACTGTCGACCTCTGCTGTAACCGGCCCCCTGATGCTTGTGAGGACCACTGAAT 829
Db      233 CTTCAGAACTGTCGACCTCTGCTGTAACCGGCCCCCTGATGCTTGTGAGGACCACTGAAT 174
Qy      830 ATTCCTCTGAGCTGACATATGAGGAGTGCAGAGCTGATCTTTATTTGAATGTTCCAGGGTC 889
Db      173 ATTCCTCTGAGCTGACATATGAGGAGTGCAGAGCTGATCTTTATTTGAATGTTCCAGGGTC 114
Qy      890 AACCTTTGTTTCTGCGGGGTTTCCAAATCTTGAACAGCTGAGAAAATCTGGAGAGTGC 949
Db      113 AACCTTTGTTTCTGCGGGGTTTCCAAATCTTGAACAGCTGAGAAAATCTGGAGAGTGC 54
Qy      950 TGGAGTCCCTTACAGAGATCTTGGCCGGAGTCTCAAGCTACCTAATCAAT 1005
Db      53 TGGAGTCCCTTACAGAGATCTTGGCCGGAGTCTCAAGCTACCTAATCAAT 1

```

```

RESULT 11
CD636757      602 bp      mRNA      linear      EST 12-JAN-2004
LOCUS         CD636757
DEFINITION    CD636757 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION     CD636757.1 GI:40285024
VERSION       CD636757.1
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens

```

```

REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE          1 (bases 1 to 602)
JOURNAL        Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
COMMENT        Circular rapid amplification of cDNA ends for high-throughput
                extension cloning of partial genes
                Genomics 84 (1), 205-210 (2004)
                Contact: Fu GK
                Incyte Genomics, Inc.
                3160 Porter Dr., Palo Alto, CA 94304, USA
                Tel: 6508454102
                Email: gfu@incyte.com

```

```

FEATURES
    source
        1..602
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone_1ib="FLP"
            /note="Vector: pDrive Cloning Vector"

```

ORIGIN

```

Query Match      43.3%; Score 566.6; DB 6; Length 602;
Best Local Similarity 99.0%; Pred. No. 7.4e-160;
Matches 591; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
Qy      410 ATGCAGAGAGAGAGTCCCATTTACAGCTATCCGAGACCTTCTCTCGAAGGGTTTGA 469
Db      6 AGCGAGAGAGAGAGTCCCATTTACAGCTATCCGAGACCTTCTCTCGAAGGGTTTGA 65
Qy      470 AACATGCCAATATTTGCTCTCTGTCATGACATATCAACCAAGAGACATGACATTGG 529
Db      66 AACATGCCAATATTTGCTCTCTGTCATGACATATCAACCAAGAGACATGACATTGG 125
Qy      530 TTTTGAATATCATGACACAGACCTGGCCAGATATATGCTCAGATCCAGAGAGGGCTTC 589
Db      126 TTTTGAATATCATGACACAGACCTGGCCAGATATATGCTCAGATCCAGAGAGGGCTTC 185

```

```

Qy      590 ATCTCATATATGTCAGACCTTTTCATGTTTCACTTTTGCGGGGCTGGGCTACATCCACC 649
Db      186 ATCTCATATATGTCAGACCTTTTCATGTTTCACTTTTGCGGGGCTGGGCTACATCCACC 245
Qy      650 ACCAACACGTTTCTTCAAGAGGACCTGAAAACCTGAACTTACTCATCAGTCACTGGAG 709
Db      246 ACCAACACGTTTCTTCAAGAGGACCTGAAAACCTGAACTTACTCATCAGTCACTGGAG 305
Qy      710 AGCTCAAACTGGCTGATTTTGTCTTGGTCCCGGGCCAAAGTCCATTTCCAGCAACATTA 769
Db      306 AGCTCAAACTGGCTGATTTTGTCTTGGTCCCGGGCCAAAGTCCATTTCCAGCAACATTA 364
Qy      770 CTTCAGAACTGTCGACCTCTGCTGTAACCGGCCCCCTGATGCTTGTGAGGACCACTGAAT 829
Db      365 CTTCAGAACTGTCGACCTCTGCTGTAACCGGCCCCCTGATGCTTGTGAGGACCACTGAAT 424
Qy      830 ATTCCTCTGAGCTGACATATGAGGAGTGCAGAGCTGATCTTTATTTGAATGTTCCAGGGTC 889
Db      425 ATTCCTCTGAGCTGACATATGAGGAGTGCAGAGCTGATCTTTATTTGAATGTTCCAGGGTC 484
Qy      890 AACCTTTGTTTCTGCGGGGTTTCCAAATCTTGAACAGCTGAG-AAAATCTGGAGAGTGC 948
Db      485 AACCTTTGTTTCTGCGGGGTTTCCAAATCTTGAACAGCTGAGAAAATCTGGAGAGTGC 544
Qy      949 CTGGAGTCCCTTACAGAGATCTTGGCCGGAGTCTCAAGCTACCTAATCAAT 1005
Db      545 CTGGAGTCCCTTACAGAGATCTTGGCCGGAGTCTCAAGCTACCTAATCAAT 601

```

```

RESULT 12
CD636759/c    602 bp      mRNA      linear      EST 12-JAN-2004
LOCUS         CD636759/c
DEFINITION    CD636759 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION     CD636759
VERSION       CD636759.1 GI:40285026
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens

```

```

REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE          1 (bases 1 to 602)
JOURNAL        Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
COMMENT        Circular rapid amplification of cDNA ends for high-throughput
                extension cloning of partial genes
                Genomics 84 (1), 205-210 (2004)
                Contact: Fu GK
                Incyte Genomics, Inc.
                3160 Porter Dr., Palo Alto, CA 94304, USA
                Tel: 6508454102
                Email: gfu@incyte.com

```

```

FEATURES
    source
        1..602
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone_1ib="FLP"
            /note="Vector: pDrive Cloning Vector"

```

ORIGIN

```

Query Match      43.3%; Score 565.8; DB 6; Length 602;
Best Local Similarity 97.8%; Pred. No. 1.3e-159;
Matches 584; Conservative 0; Mismatches 12; Indels 1; Gaps 1;
Qy      410 ATGCAGAGAGAGAGTCCCATTTACAGCTATCCGAGAGCTTCTCTCGAAGGGTTTGA 469
Db      597 AGCGAGAGAGAGAGTCCCATTTACAGCTATCCGAGAGCTTCTCTCGAAGGGTTTGA 538
Qy      470 AACATGCCAATATTTGCTCTCTGTCATGACATATCAACCAAGAGACATGACATTGG 529
Db      537 AACATGCCAATATTTGCTCTCTGTCATGACATATCAACCAAGAGACATGACATTGG 478
Qy      530 TTTTGAATATCATGACACAGACCTGGCCAGATATATGCTCAGATCCAGAGAGGGCTTC 589

```

Db	477	TTTTGAATACATCAACAGACCTGGCCAGTATATGTCTCAGATCCAGAGGGCTTC	418
QY	590	ATCTCATATATGTGAGACTTTTCAATGTTTCAACTTTTGGCGGGGCTGGCGTACATCCACC	649
Db	417	ATCTCATATATGTGAGACTTTTCAATGTTTCAACTTTTGGCGGGGCTGGCGTACATCCACC	358
QY	650	ACCAACACGTTCTTACAGGGAGCTGAAACCTCAGAACTTACTATCATGTAACCTGGGAG	709
Db	357	ACCAACACGTTCTTACAGGGAGCTGAAACCTCAGAACTTACTATCATGTAACCTGGGAG	298
QY	710	AGCTCAACTGGGCTGATTTTGTGCTGGCCGGGGCAAGTCCATTTCCAGCCAGACATCACT	769
Db	297	AGCTCAACTGGGCTGATTTTGTGCTGGCCGGGGCAAGTCCATTTCCAGCCAGACATCACT	238
QY	770	CTTCAGAAAGTGTACCCCTCTGTGATACCGGCCCCCTGATAGCTTTTGTCTGGAGACCACTGAT	829
Db	237	CTTCAGAAAGTGTACCCCTCTGTGATACCGGCCCCCTGATAGCTTTTGTCTGGAGACCACTGAT	178
QY	830	ATTCTCTGAGCTGGACATATAGGGGTGACAGGCTGATCTTTATTTGAAATGTTCCAGGGCT	889
Db	177	ATTCTCTGAGCTGGACATATAGGGGTGACAGGCTGATCTTTATTTGAAATGTTCCAGGGCT	118
QY	890	AACCTTTGTTCTGGGGGTTTCCAAAT - CTTTGAACAGCTGGAGAAATCTGGAGGTG	948
Db	117	AACCTTTGTTCTGGGGGTTTCCAAAT - CTTTGAACAGCTGGAGAAATCTGGAGAGGTG	58
QY	949	CTGGGAGTCCCTACAGAGGATATCTTGGCCGGGAGTCTCCAGCTACCTAACTACAT	1005
Db	57	CTGGGAGTCCCTACAGAGGATATCTTGGCCGGGAGTCTCCAGCTACCTAACTACAT	1

LOCUS	CD636746/c	601 bp	mRNA	linear	EST 12-JAN-2004
DEFINITION	5601492501	FLP Homo sapiens cDNA, mRNA sequence.			
ACCESSION	CD636746				
VERSION	CD636746.1	GI:40285013			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 601)				
AUTHORS	Fu,G.K., Wang,J.T., Yang,T., Au-Young,J. and Stuve,L.L.				
TITLE	Circular rapid amplification of cDNA ends for high-throughput extension cloning of partial genes				
JOURNAL	Genomics 84 (1), 205-210 (2004)				
COMMENT	Contact: Fu GK Incyte Genomics, Inc. 3160 Porter Dr., Palo Alto, CA 94304, USA Tel: 6508454102 Email: gfu@incyte.com.				
FEATURES	Location/Qualifiers				
Source	1..601 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone_lib="FLP" /note="Vector: pDrive Cloning Vector"				
ORIGIN					
Query Match	43.1%; Score 564.2; DB 6; Length 601;				
Best Local Similarity	99.2%; Pred. No.3.9e-159;				
Matches	588; Conservative 0; Mismatches 3; Indels 2; Gaps 2;				
QY	412 GCAAGAGGAAGAGTCCCATTTACAGCTATCCGAGAAGCTTCTCTCTGAAGGGTTTGA	471			
Db	593 GCAAGAGGAAGAGTCCCATTTACAGCTATCCGAGAAGCTTCTCTCTGAAGGGTTTGA	534			
QY	472 CATGCCAATATTGCGTCTCGATGATCACTAATATCCACACAAAGACACTGACATTGCTT	531			
Db	533 CATGCCAATATTGCGTCTCGATGATCACTAATATCCACACC-AAGAACACTGACATTGCTT	475			

OY	532	TTTGAAATACATGCAACAAGACCTGGGCCAGTATATGTCTCAGCATCCAGAGGGCTTCAT	591
Db	474	TTTGAATACAGCAACAAGACCTGGGCCAGTATATGTCTCAGCATCCAGAGGGCTTCAT	415
OY	592	CCTGATATATGCAACATCTTTCATCTTTCAACTTTTGCGGGGCTGAGGCTACATCCACAC	651
Db	414	CCTGATATATGCAACATCTTTCATCTTTCAACTTTTGCGGGGCTGAGGCTACATCCACAC	355
OY	652	CAACACGTTCTTCAACAGGAGCTTGAAACCTTCAGAACTTATCTCATCAGTCACTCTGGGAG	711
Db	354	CAACACGTTCTTCAACAGGAGCTTGAAACCTTCAGAACTTATCTCATCAGTCACTCTGGGAG	295
OY	712	CTCAAACTGGCTGATTTTGGCTTGGCTTGGCCGGGSCAAGTCCATTTCCAGCACAACATACTCT	771
Db	294	CTCAAACTGGCTGATTTTGGCTTGGCTTGGCCGGGSCAAGTCCATTTCCAGCACAACATACTCT	235
OY	772	TCAGAACTCGTAGCCCTCTGTAACCGGCCCCCTGATGCTTTGCTGGGAGCACTGAAATAT	831
Db	234	TCAGAACTCGTAGCCCTCTGTAACCGGCCCCCTGATGCTTTGCTGGGAGCACTGAAATAT	175
OY	832	TCCCTCTAGCTGAGCAATATGGGGGGGAGAGGCTGCAGACTTTATTTGAATGTTCCAGGGTTAA	891
Db	174	TCCCTCTAGCTGAGCAATATGGGGGGGAGAGGCTGCAGACTTTATTTGAATGTTCCAGGGTTAA	115
OY	892	CCTTTGTGTTCTCGGGGTTTCCAAATCTCTTGAAACAGCTGAGAAATCTGSGAGTGCTG	951
Db	114	CCTTTGTGTTCTCGGGGTTTCCAAATCTCTTGAAACAGCTGAGAAATCTGSGAGTGCTG	55
OY	952	CGAGTCCCTACAGAGGATCTTGG-CCGGAGTCTCCAAAGTACTCTAACCTTCA 1003	
Db	54	GGAGTCCCTACAGAGGATCTTGNCCAGGGGCTCTCAAAGTACTCTAACCTTCA 2	

CD636755	RESULT 14
LOCUS	CD636755
DEFINITION	56015017H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION	CD636755
VERSION	CD636755.1 GI:40285022
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 602) Fu,G.K., Wang,J.T., Yang,D., Au-Young,J., and Stuve,L.L. Circular rapid amplification of cDNA ends for high-throughput extension cloning of partial genes
AUTHORS	Genomics 84 (1), 205-210 (2004)
TITLE	Contact: Fu GK
JOURNAL	Incyte Genomics, Inc. 3160 Porter Dr., Palo Alto, CA 94304, USA Tel.: 6509454102
COMMENT	Email: gfu@incyte.com.
FEATURES	Location/Qualifiers
source	1..602
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone_lib="FLP"
	/note="Vector: pDrive Cloning Vector"
ORIGIN	
Query Match	42.8%; Score 560; DB 6; Length 602;
Beet Local Similarity	97.3%; Pred. No. 7.4e-158;
Matches 580; Conservative	0; Mismatches 15; Indels 1; Gaps 11
Qy	410 ATGCAGAGAAGAGATCCATTTCACGCTATCGAAGACTTCTCTCCTGAAGGGTTTGA 469
Dd	6 AGGCGAGAGAAAGAGTCCCATTTACGTATCCGAGAACTTCTCTCCTGAAGGGTTTGA 65
Qy	470 AACATGCCAATATTGTGCTCCTCGCATGACATTAATTCACACCAAGAGACA CTGACATTGC 529

```

Db      66 AACATGCCAATCTGTGCTCTGCATGACATTAATCCACACCAAGAGACATGCACTTGG 125
Qy      530 TTTTGAATACATGACACAGACCTGGCCAGATATATGTCTCAGATCCAGAGAGGGCTTC 589
Db      126 TTTTGAATACATGACACAGACCTGGCCAGATATATGTCTCAGATCCAGAGAGGGCTTC 185
Qy      530 ATCTCATATATGTCAGACTTTTTCATTTTCACTTTTGGGGGCGCTGGGCTACATCACC 649
Db      186 ATCTCATATATGTCAGACTTTTTCATTTTCACTTTTGGGGGCGCTGGGCTACATCACC 245
Qy      650 ACCAACACGTTCTTCAACAGGAGCCTGAAACCTCAGAACTTACTCATCAGATCACTGGAG 709
Db      246 ACCAACACGTTCTTCAACAGGAGCCTGAAACCTCAGAACTTACTCATCAGATCACTGGAG 305
Qy      710 AGCTCAAACTGCTGATTTTGGTCTTGGCCCGGCGCAAGTCATTTCCACGACAGACATCT 769
Db      306 AGCTCAAACTGCTGATTTTGGTCTTGGCCCGGCGCAAGTCATTTCCACGACAGACATCT 365
Qy      770 CTTCAGAACTGCTGACCTCTGCTGACCGGCCCCCTGATGCTTTGCTGGAGCCACTGAAT 829
Db      366 CTTCAGAACTGCTGACCTCTGCTGACCGGCCCCCTGATGCTTTGCTGGAGCCACTGAAT 425
Qy      830 ATTCCTGAGCTGACATATGAGGGGTCAGGCTGACATTTATGAAAGTCCAGGGCTC 889
Db      426 ATTCCTGAGCTGACATATGAGGGGTCAGGCTGACATTTATGAAAGTCCAGGGCTC 484
Qy      890 AACCTTTGTTTCTGCGGGGTTTCCACATCTTGAACAGCTGAGAAATCTGGAGAGTGC 949
Db      485 AACCTTTGTTTCTGCGGGGTTTCCACATCTTGAACAGCTGAGAAATCTGGAGAGTGC 544
Qy      950 TGGGAGTCCCTACAGAGATCTTGGCCGGAGCTCTCAAGCTAACTTAACAAT 1005
Db      545 TGGGAGTCCCTACAGAGATCTTGGCCGGAGCTCTCAAGCTAACTTAACAAT 600

```

RESULT 15

```

CD636750/c  CD636750 587 bp mRNA linear EST 12-JAN-2004
LOCUS        56014981J1 FLP Homo sapiens cDNA, mRNA sequence.
DEFINITION   CD636750
ACCESSION    CD636750.1 GI:40285017
VERSION      EST.
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens

```

REFERENCE

```

AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE        1 (bases 1 to 587)
JOURNAL      Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
COMMENT      Circular rapid amplification of cDNA ends for high-throughput
              Genomics 84 (1), 205-210 (2004)

```

JOURNAL

```

COMMENT      Contact: Fu GK
              Incyte Genomics, Inc.
              3160 Porter Dr., Palo Alto, CA 94304, USA
              Tel: 6508454102
              Email: gfu@incyte.com.

```

FEATURES

```

source       1..587
              Location/Qualifiers
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone_1ib="FLP"
              /note="Vector: pDrive Cloning Vector"

```

ORIGIN

```

Query Match 42.6%; Score 557.2; DB 6; Length 587;
Best Local Similarity 99.3%; Pred. No. 5.2e-157;
Matches 570; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
424 GTCCATTACAGATATCCGAGAGCTTCTCTGAGAGGTTTGAACATGCAATATT 483
Db      580 GTCCATTACAGATATCCGAGAGCTTCTCTGAGAGGTTTGAACATGCAATATT 521.

```

```

Qy      484 GTCTCTGATGACATTAATCCACACCAAGAGACATGCAATGCTTTTGAATCATG 543
Db      520 GTCTCTGATGACATTAATCCACACCAAGAGACATGCAATGCTTTTGAATCATG 461
Qy      544 CACACAGACCTTGGCCAGATATATGTCTCAGATCCAGAGAGGGCTTCACTCAATATGTC 603
Db      460 CACACAGACCTTGGCCAGATATATGTCTCAGATCCAGAGAGGGCTTCACTCAATATGTC 401
Qy      604 AGACTTTTCATGTTTCAACTTTTGGGGGCTGGGGTATCATCACACCAACACGTTCTT 663
Db      400 AGACTTTTCATGTTTCAACTTTTGGGGGCTGGGGTATCATCACACCAACACGTTCTT 341
Qy      664 CACAGGACCTGAACCTCAGAACTTACTCATCAGTCACTGGAGAGCTCAAACTGGCT 723
Db      340 CACAGGACCTGAACCTCAGAACTTACTCATCAGTCACTGGAGAGCTCAAACTGGCT 281
Qy      724 GATTTTGGTCTTGGCCCGGCGCAAGTCATTTCCAGCGAGACATCTTCAAGAGTGTG 783
Db      280 GATTTTGGTCTTGGCCCGGCGCAAGTCATTTCCAGCGAGACATCTTCAAGAGTGTG 221
Qy      784 ACCCTCTGATACCGGCCCCCTGATGCTTTGCTGGAGCCACTGAATATTCCTGAGCTG 843
Db      220 ACCCTCTGATACCGGCCCCCTGATGCTTTGCTGGAGCCACTGAATATTCCTGAGCTG 161
Qy      844 GACATATGAGGGTGCAGGCTGCATCTTTATGAAATGTTCCAGGGTCAACCTTTGTTCT 903
Db      160 GACATATGAGGGTGCAGGCTGCATCTTTATGAAATGTTCCAGGGTCAACCTTTGTTCT 101
Qy      904 GGGGTTTCCAAATCCTTGAACAGCTGAGAGAAATCTGGAGAGTCTGGAGATCCCTACA 963
Db      100 GGGGTTTCCAAATCCTTGAACAGCTGAGAGAAATCTGGAGAGTCTGGAGATCCCTACA 41
Qy      964 GAGATATCTTGGCCGGAGTCTCAAGCTACCTA 997
Db      40 GAGATATCTT-GCCGGAGTCTCCAGCTACCAA 8

```

Search completed: December 27, 2004, 10:56:52
Job time : 2505.82 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2004, 17:36:56 ; Search time 377.27 Seconds
(without alignment)
18199.836 Million cell updates/sec

Title: US-10-786-065-4

Perfect score: 1308
Sequence: 1 atgggtcaagagcgtctgtgc.....agtttagcaaatcgtctgtga 1308

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1908s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1308	100.0	1308	6	AA148891	Aa148891 Human pfc
2	1308	100.0	1308	6	ACC79970	Acc79970 Human ser
3	1308	100.0	1584	9	ACC79968	Acc79968 Human ser
4	1308	100.0	1628	9	ACC79971	Acc79971 Human ser
5	1308	100.0	2250	12	ADJ96563	Adj96563 Human cyc
6	1191	91.1	2140	11	ADL22554	Adl22554 Human dis
7	1186	90.7	1311	10	ADC30755	Adc30755 Human nov
8	1185.4	90.6	1534	9	ACC79969	Acc79969 Human ser
9	1185.4	90.6	1534	9	ACC79967	Acc79967 Human ser
10	1009	77.1	2203	6	AA148889	Aa148889 Human pfc
11	887.6	67.9	1376	8	ABA00722	AbA00722 Human krp
12	873	66.7	1828	12	ADM16422	Adm16422 DNA encod
13	821	62.8	888	10	ABE77160	AbE77160 Human pro
14	702.2	53.7	1077	4	AA144668	Aa144668 Novel pro
15	702.2	53.7	1077	12	AD129366	Ad129366 Human MAR
16	481	36.8	481	6	AA17061	Aa17061 Human pro
17	481	36.8	481	6	AA17061	Aa17061 Human pro
18	481	36.8	481	6	AA17061	Aa17061 Human pro
19	439.6	33.6	4529	12	ABX14985	Abx14985 Marker ge
20	437.8	33.5	1272	2	AA186605	Aa186605 Coding re
21	436.4	33.4	2424	2	AA186606	Aa186606 Murine mp

22	436.2	33.3	4957	10	ADD99960	Add99960 Human can
23	436.2	33.3	4957	11	ADN95587	Adn95587 Human BRC
24	436.2	33.3	4957	12	ADJ75096	Adj75096 Marker ge
25	436.2	33.3	4957	12	ADQ17589	Adq17589 Human sof
26	436.2	33.3	5161	12	ADO22276	Ado22276 Human sof
27	434.6	33.2	1410	6	AA153644	Aa153644 DNA enco
28	329.2	25.2	2703	4	AB120823	Ab120823 Drosophill
29	329.2	25.2	2821	4	AB120823	Ab120823 Drosophill
30	329.2	25.2	3373	4	AB120821	Ab120821 Drosophill
31	299.8	22.9	3024	4	AA199648	Aa199648 Human pro
32	297.8	22.8	1745	6	ABK84625	Abk84625 Human CDN
33	297.8	22.8	1745	6	AA161691	Aa161691 Human PCT
34	297.8	22.8	1745	10	ADK38402	Adk38402 Human pro
35	297.8	22.8	1745	10	ADK61358	Adk61358 Ovarian C
36	297.8	22.8	2382	12	ADM86949	Adm86949 Human pro
37	297.8	22.8	2987	9	AA161777	Aa161777 Human PCT
38	297.8	22.8	3131	12	ADQ19443	Adq19443 Human sof
39	297.8	22.8	3182	9	AA161776	Aa161776 Human PCT
40	297.4	22.7	2825	3	AA121685	Aa121685 Human bre
41	296.6	22.7	1839	3	AA146142	Aa146142 CDNA sequ
42	296.6	22.7	2580	4	AA158117	Aa158117 Human pol
43	296.6	22.7	2858	4	AA188828	Aa188828 Human kin
44	296.2	22.6	2363	4	AAK51826	AaK51826 Human pol
45	296.2	22.6	2868	4	AA158177	Aa158177 Human pol

ALIGNMENTS

RESULT 1	AA148891	strandard; cDNA, 1308 BP.
ID	AA148891	strandard; cDNA, 1308 BP.
XX	AA148891;	
AC	AA148891;	
XX		
DT	24-OCT-2002	(first entry)
XX		
DE	Human pfc	family kinase splice form 2 coding sequence.
XX		
KW	Human; pfc	family kinase; kinase; enzyme; testis; brain; cytosolic;
KW	uterus endometrium	adenocarcinoma; lung fibroblast; splice form;
KW	kidney renal cell	adenocarcinoma; gene therapy; gene; ss..
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1306
FT		/*tag= a
FT		/product= "kinase"
XX		
PN	WO200261060-A2.	
XX		
PD	08-AUG-2002.	
XX		
PF	17-JAN-2002;	2002WO-US001106.
XX		
PR	31-JAN-2001;	2001US-0265151P.
PR	09-MAR-2001;	2001US-00801861.
XX		
PA	(PEKE)	PE CORP NY.
XX		
PI	Yan C, Ketchum K, Di	Francesco V, Beasley EM;
XX		
DR	WPI; 2002-608515/65.	
DR	P-PSDB; AAO18614.	
XX		
PT	New human kinase	peptide and nucleic acid molecule, useful for treating
PT	disorders associated	with abnormal expression of kinase protein, e.g.
PT	adenocarcinoma of	uterus or lung, in drug screening assays and
PT	pharmacogenomic	analysis.
XX		
PS	Claim 4; Fig 1; 131p;	English.
XX		

CC The present invention provides the protein, cDNA and gene sequences of
 CC two splice variants of a human Pfkfb1 family kinase. The sequences are
 CC specifically expressed in the human testis, brain, uterus endometrium
 CC adenocarcinomas, lung fibroblasts, kidney renal cell adenocarcinomas, and
 CC can be used to treat related diseases. The present sequence is the cDNA
 CC of splice variant 2 of the invention
 XX

Sequence 1308 BP, 328 A; 332 C; 330 G; 318 T; 0 U; 0 Other;

Query Match 100.0%; Score 1308; DB 6; Length 1308;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGGGTCAAGAGCTGTGTGCAAGAGCTGTACAGCTGTGATGCGAGCTGTCTTCCATTGTTCA 60
DB 1 ATGGGTCAAGAGCTGTGTGCAAGAGCTGTACAGCTGTGATGCGAGCTGTCTTCCATTGTTCA 60
QY 61 GAGGAGGCGAGGACACAGCTGTGAGGAGTCAAGCTGTGAGACCAAGAGGCTGTGCTTC 120
DB 61 GAGGAGGCGAGGACACAGCTGTGAGGAGTCAAGCTGTGAGACCAAGAGGCTGTGCTTC 120
QY 121 AAGCTTAACAGACCTTAAAGAGCATCATGTTCCATTTACCCCAAGGAGACTT 180
DB 121 AAGCTTAACAGACCTTAAAGAGCATCATGTTCCATTTACCCCAAGGAGACTT 180
QY 181 CAAGCTGCCCCGTCGCAAGTTCAGAGTAAAGGCGCAGAGTAAAGTATGTTT 240
DB 181 CAAGCTGCCCCGTCGCAAGTTCAGAGTAAAGGCGCAGAGTAAAGTATGTTT 240
QY 241 CAGGAAGAGATCTGAGGAGGCTTTTCAAGTGAAGAGAGCTCCCTTTTGGGAGGCG 300
DB 241 CAGGAAGAGATCTGAGGAGGCTTTTCAAGTGAAGAGAGCTCCCTTTTGGGAGGCG 300
QY 301 TCATCTTACTTGAACCTTGGAAAGCTGGGTGAAGGCTCTTATGCGACAGTTTCAAAGGG 360
DB 301 TCATCTTACTTGAACCTTGGAAAGCTGGGTGAAGGCTCTTATGCGACAGTTTCAAAGGG 360
QY 361 ATTAGCAGAAATAAATGAGCAACTAGTGGCTTTAAAGTATCAAGCATGAAATGAGAGGA 420
DB 361 ATTAGCAGAAATAAATGAGCAACTAGTGGCTTTAAAGTATCAAGCATGAAATGAGAGGA 420
QY 421 GGAAGTCCATTTACAGCTATCCAGAGAGCTTCTCTGTAAGGGTTTGAAGTGCAT 480
DB 421 GGAAGTCCATTTACAGCTATCCAGAGAGCTTCTCTGTAAGGGTTTGAAGTGCAT 480
QY 481 ATTGTCTCTGATGATCATATCCACACCAAGAGACCTGACATTCGTTTGAATAC 540
DB 481 ATTGTCTCTGATGATCATATCCACACCAAGAGACCTGACATTCGTTTGAATAC 540
QY 541 ATGCACACAGACCTGAGCCAGTATATGTCACAGATCCAGAGGGCTTCACTCTCAAT 600
DB 541 ATGCACACAGACCTGAGCCAGTATATGTCACAGATCCAGAGGGCTTCACTCTCAAT 600
QY 601 GTTCAGACTTTTCAATGTTTCACTTTTGCAGGGCTGTGCTGTACATCCACCAACACGTT 660
DB 601 GTTCAGACTTTTCAATGTTTCACTTTTGCAGGGCTGTGCTGTACATCCACCAACACGTT 660
QY 661 CTTCACAGGAGCTGTGAACCTCAGAACTTATCATGATGTCACCTGGAGAGGCTCAAACTG 720
DB 661 CTTCACAGGAGCTGTGAACCTCAGAACTTATCATGATGTCACCTGGAGAGGCTCAAACTG 720
QY 721 GGTGATTTTGTCTTGTGCGGGGCAAGTCAATTCACAGCAGACATACCTTCCAGAGT 780
DB 721 GGTGATTTTGTCTTGTGCGGGGCAAGTCAATTCACAGCAGACATACCTTCCAGAGT 780
QY 781 GTGACCTCTGTGATCGGCGCCCTGTATGCTTTGCTGGAGGACATGATATATCTCTGAG 840
DB 781 GTGACCTCTGTGATCGGCGCCCTGTATGCTTTGCTGGAGGACATGATATATCTCTGAG 840
QY 841 CTGGACATATGGGGGCTGAGCTGATCTTTATGAAATGTTCAAGGCTCAACCTTTGTTT 900
DB 841 CTGGACATATGGGGGCTGAGCTGATCTTTATGAAATGTTCAAGGCTCAACCTTTGTTT 900

```

```

QY 901 CTGGGGGTTTCCAACTCTTTGAAAGCTGAGGAAATCTGGAGGTGCTGGAGTCCCT 960
DB 901 CTGGGGGTTTCCAACTCTTTGAAAGCTGAGGAAATCTGGAGGTGCTGGAGTCCCT 960
QY 961 ACAGAGGATACCTGGCCGGGAGTCTCCAGACTACCTAATCAATCCAGAAATGTTCCCA 1020
DB 961 ACAGAGGATACCTGGCCGGGAGTCTCCAGACTACCTAATCAATCCAGAAATGTTCCCA 1020
QY 1021 CTGCTTACGCTCGAAGCTTCTGATGTTCTGGAACAGGCTGGGAGGTTCTTGAAGCT 1080
DB 1021 CTGCTTACGCTCGAAGCTTCTGATGTTCTGGAACAGGCTGGGAGGTTCTTGAAGCT 1080
QY 1081 GAAAGACTGCGCTCCCAATGCTTAAAGGCTTTCCCAAGACCGGCTCTCCGCCAGGAA 1140
DB 1081 GAAAGACTGCGCTCCCAATGCTTAAAGGCTTTCCCAAGACCGGCTCTCCGCCAGGAA 1140
QY 1141 GCACCTTGTTCATGATTAATTTTACGCGCCCTGACCATCTACAGCTTACAGCTTCTGATGAG 1200
DB 1141 GCACCTTGTTCATGATTAATTTTACGCGCCCTGACCATCTACAGCTTACAGCTTCTGATGAG 1200
QY 1201 GAGTCTTTGTTTACAGTTTCAAGAGTGAAGCTTAAAGCCAGAAATGTGACCTTTTGGCC 1260
DB 1201 GAGTCTTTGTTTACAGTTTCAAGAGTGAAGGCTTAAAGCCAGAAATGTGACCTTTTGGCC 1260
QY 1261 TCTTACAGAAAGGTCAACCAACCCAGCTTACCAATGTGTGTA 1308
DB 1261 TCTTACAGAAAGGTCAACCAACCCAGCTTACCAATGTGTGTA 1308

```

RESULT 2

ACC79970
 ID ACC79970 Blandard; cDNA, 1308 BP.

AC79970;
 DT 09-SEP-2003 (first entry)

XX
 DE Human serine/threonine protein kinase encoding cDNA SEQ ID NO:8.

XX Human; serine/threonine protein kinase; kinase; enzyme; cytosolic;
 KW antidiabetic; neuroprotective; nootropic; antiparkinsonian; analgesic;
 KW cerebroprotective; antiinflammatory; gastrointestinal; hepatotropic;
 KW gynecological; uropathic; dermatological; gene therapy; cancer;
 KW diabetes; central nervous system disorder; CNS disorder; liver disease;
 KW respiratory disorder; chronic obstructive pulmonary disease; stroke;
 KW cardiovascular disorder; dermatological disorder; urological disorder;
 KW gastrointestinal disease; haematological disorder; Alzheimer's disease;
 KW musculoskeletal disorder; reproductive disorder; Parkinson's disease;
 KW neuropathic pain; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH 1. 1308
 FT CDS /tag= a
 FT /product= "serine/threonine protein kinase"

XX WO2003046167-A1.

XX 05-JUN-2003.

XX 26-NOV-2002; 2002MO-EP013268.

XX 27-NOV-2001; 2001US-0333131P.

XX (FARB) BAYER AG.

XX Koehler RH;

XX WPI; 2003-505196/47.

XX DR P-PSDB; ABR57361.

XX New polynucleotide encoding a serine/threonine protein kinase

PT polypeptide, useful for diagnosing, preventing or treating diseases
 associated with serine/threonine protein kinase dysfunction, e.g. cancer
 or diabetes.

PS Disclosure; Fig 8; 196bp; English.

XX The present sequence encodes a human serine/threonine protein kinase (I).
 CC (I) has cytosolic, antidiabetic, neuroprotective, antiparkinsonian,
 CC neurotrophic, cerebroprotective, analgesic, antiinflammatory, hepatotropic,
 CC gastrointestinal, gynaecological, uropathic and dermatological
 CC activities, and can be used in gene therapy. Serine/threonine protein
 CC kinase polynucleotide and polypeptide sequences can be used in
 CC diagnosing, preventing, ameliorating or treating diseases associated with
 CC serine/threonine protein kinase dysfunction. They may also be used to
 CC identify test compounds that may act, for example, as activators or
 CC inhibitors at the enzyme's active site. The human serine/threonine
 CC protein kinase and its fragments are also useful in raising specific
 CC antibodies that can block the enzyme and effectively reduce its activity.
 CC Human serine/threonine protein kinase sequences can be used in the
 CC preparation of a medicament for modulating the activity of a serine/
 CC threonine protein kinase in a disease, e.g. cancer, diabetes, a central
 CC nervous system (CNS) disorder, a respiratory disorder (including chronic
 CC obstructive pulmonary disease), a cardiovascular disorder, a
 CC dermatological disorder, a gastrointestinal or liver disease, a
 CC haematological disorder, a musculoskeletal disorder, a reproductive
 CC disorder, or a urological disorder. CNS disorders may include Alzheimer's
 CC disease, Parkinson's disease, stroke or neuropathic pain

CC Sequence 1308 BP; 328 A; 332 C; 330 G; 318 T; 0 U; 0 Other;

Query Match 100.0%; Score 1308; DB 9; Length 1308;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGGTCAGAGAGCTGTGTCAAGAGCTGTACAGCTGTGATGCTACTACATTGTTCA 60
 1 ATGGGTCAGAGAGCTGTGTCAAGAGCTGTACAGCTGTGATGCTACTACATTGTTCA 60
 61 GAGGAGGCGGAGGACACAGCTGTGTGAGAGAGTCACTGTAGACACAGAGGCTGCTTC 120
 61 GAGGAGGCGGAGGACACAGCTGTGTGAGAGAGTCACTGTAGACACAGAGGCTGCTTC 120
 61 GAGGAGGCGGAGGACACAGCTGTGTGAGAGAGTCACTGTAGACACAGAGGCTGCTTC 120
 121 AAGCTTAACAGAGCTTAAGAGAGTCACTGTATCTTCAATCTTCAATCTTCAATCTTCA 180
 121 AAGCTTAACAGAGCTTAAGAGAGTCACTGTATCTTCAATCTTCAATCTTCAATCTTCA 180
 121 AAGCTTAACAGAGCTTAAGAGAGTCACTGTATCTTCAATCTTCAATCTTCAATCTTCA 180
 181 CAGCTGCGGCTGCGGAGAGTCAAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAG 240
 181 CAGCTGCGGCTGCGGAGAGTCAAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAG 240
 181 CAGCTGCGGCTGCGGAGAGTCAAGAGTCAAGAGAGTCAAGAGAGTCAAGAGAGTCAAG 240
 241 CAGGAGAGAGATCTGAGGAGGCTTTCAAGTGAAGAGAGGCTCCCTTTTGGGAGGCC 300
 241 CAGGAGAGAGATCTGAGGAGGCTTTCAAGTGAAGAGAGGCTCCCTTTTGGGAGGCC 300
 241 CAGGAGAGAGATCTGAGGAGGCTTTCAAGTGAAGAGAGGCTCCCTTTTGGGAGGCC 300
 301 TCATCTTACTTGAACCTTGAAGAGCTGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTT 360
 301 TCATCTTACTTGAACCTTGAAGAGCTGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTT 360
 301 TCATCTTACTTGAACCTTGAAGAGCTGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTT 360
 361 ATTAGCAGATTAATAGCAGACAGTGAAGGCTTTAAAGTATGATGATGATGATGATGAT 420
 361 ATTAGCAGATTAATAGCAGACAGTGAAGGCTTTAAAGTATGATGATGATGATGATGAT 420
 361 ATTAGCAGATTAATAGCAGACAGTGAAGGCTTTAAAGTATGATGATGATGATGATGAT 420
 421 GAGGTCCTTGAACCTTGAAGAGCTGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGA 480
 421 GAGGTCCTTGAACCTTGAAGAGCTGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGA 480
 421 GAGGTCCTTGAACCTTGAAGAGCTGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGA 480
 481 ATTGTCCTTGAACCTTGAAGAGCTGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGA 540
 481 ATTGTCCTTGAACCTTGAAGAGCTGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGA 540
 481 ATTGTCCTTGAACCTTGAAGAGCTGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGA 540
 541 ATGCACACAGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 600
 541 ATGCACACAGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 600

QY 601 GTGACACTTTTCAATGTTTCAATCTTTTGGGAGGCTGCTGATCCACCAACACGTT 660
 DB 601 GTGACACTTTTCAATGTTTCAATCTTTTGGGAGGCTGCTGATCCACCAACACGTT 660
 QY 661 CTTTCAGAGGAGCTGTAACCTGAGAACTTCAATCAATCAATCAATCAATCAATCAAT 720
 DB 661 CTTTCAGAGGAGCTGTAACCTGAGAACTTCAATCAATCAATCAATCAATCAATCAAT 720
 QY 721 GCTGATTTTGGTCTTCCGCGGAGCAAGTCAATCTTCCAGCAGACATCTTCCAGAGTC 780
 DB 721 GCTGATTTTGGTCTTCCGCGGAGCAAGTCAATCTTCCAGCAGACATCTTCCAGAGTC 780
 QY 781 GTGACCTTGTGATCCGCGGAGCTGATGCTTGTGAGAGCACTGATATTTCTGTAG 840
 DB 781 GTGACCTTGTGATCCGCGGAGCTGATGCTTGTGAGAGCACTGATATTTCTGTAG 840
 QY 841 CTGAGATATATGAGGAGTCAAGCTGATCTTATTTGAATGTTCCAGGCTCACTTGTGTT 900
 DB 841 CTGAGATATATGAGGAGTCAAGCTGATCTTATTTGAATGTTCCAGGCTCACTTGTGTT 900
 QY 901 CCGGAGGTTTCCAAATCTTGAACAGCTGAGAAATCTGGAGAGTGTGAGTCCCT 960
 DB 901 CCGGAGGTTTCCAAATCTTGAACAGCTGAGAAATCTGGAGAGTGTGAGTCCCT 960
 QY 961 ACAGAGATATCTTGGCGGAGAGTCTCAAGCTTCACTTCACTTCACTTCACTTCACT 1020
 DB 961 ACAGAGATATCTTGGCGGAGAGTCTCAAGCTTCACTTCACTTCACTTCACTTCACT 1020
 QY 1021 CTGACCTAGGCTTCAAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1080
 DB 1021 CTGACCTAGGCTTCAAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1080
 QY 1081 GAAAGACCTGCTTCCAGATGCTTAAAGAGCTTCCAGAGAGCGCTCTCCGCCAGAA 1140
 DB 1081 GAAAGACCTGCTTCCAGATGCTTAAAGAGCTTCCAGAGAGCGCTCTCCGCCAGAA 1140
 QY 1141 GCACTTGTATGATTAATTTTCAAGCGGCTGCTGATCTCACTTCACTTCACTTCACT 1200
 DB 1141 GCACTTGTATGATTAATTTTCAAGCGGCTGCTGATCTCACTTCACTTCACTTCACT 1200
 QY 1201 GAGCTTGTATGATTAATTTTCAAGCGGCTGCTGATCTCACTTCACTTCACTTCACT 1260
 DB 1201 GAGCTTGTATGATTAATTTTCAAGCGGCTGCTGATCTCACTTCACTTCACTTCACT 1260
 QY 1261 TCTTACAGAGAGGCTCAACCAAGGCTTGAAGAGTGTGAGTGTGAGTGTGAGTGTG 1308
 DB 1261 TCTTACAGAGAGGCTCAACCAAGGCTTGAAGAGTGTGAGTGTGAGTGTGAGTGTG 1308

RESULT 3
 ACCT9968
 ID ACCT9968 standard; cDNA; 1584 BP.
 XX
 AC ACCT9968;
 DT 09-SEP-2003 (first entry)
 XX
 DE Human serine/threonine protein kinase encoding cDNA seq ID NO:4.
 XX
 KW Human; serine/threonine protein kinase; kinase; enzyme; cytosolic;
 KW antidiabetic; neuroprotective; neurotrophic; antiparkinsonian; analgesic;
 KW cerebroprotective; antiinflammatory; gastrointestinal; hepatotropic;
 KW gynaecological; uropathic; dermatological; gene therapy; cancer;
 KW diabetes; central nervous system disorder; CNS disorder; liver disease;
 KW respiratory disorder; chronic obstructive pulmonary disease; stroke;
 KW cardiovascular disorder; dermatological disorder; urological disorder;
 KW gastrointestinal disease; haematological disorder; Alzheimer's disease;
 KW musculoskeletal disorder; reproductive disorder; Parkinson's disease;
 KW neuropathic pain; gene; ss.
 XX
 OS Homo sapiens.

Key Location/Qualifiers
FT CDS 1..1308
FT /tag= a
FT /product= "serine/threonine protein kinase"
PN MO2003046167-A1.
XX 05-JUN-2003.
XX
XX 26-NOV-2002; 2002MO-EP013268.
XX
XX 27-NOV-2001; 2001US-0333131P.
XX
XX (PAB) BAYER AG.
XX
XX Koehler RH;
XX WPI; 2003-505196/47.
XX P-PSDB; ABR57359.
XX
XX
XX New polynucleotide encoding a serine/threonine protein kinase
XX polypeptide, useful for diagnosing, preventing or treating diseases
XX associated with serine/threonine protein kinase dysfunction, e.g. cancer
XX or diabetes.
XX
XX Disclosure; Fig 4; 196pp, English.
XX
XX The present sequence encodes a human serine/threonine protein kinase (1).
XX (1) has cytostatic, antidiabetic, neuroprotective, antiparkinsonian,
XX neurotropic, cerebroprotective, analgesic, antiinflammatory, hepatotropic,
XX gastrointestinal, gynaecological, uropathic and dermatological
XX activities, and can be used in gene therapy. Serine/threonine protein
XX kinase polynucleotide and polypeptide sequences can be used in
XX diagnosing, preventing, ameliorating or treating diseases associated with
XX serine/threonine protein kinase dysfunction. They may also be used to
XX identify test compounds that may act, for example, as activators or
XX inhibitors at the enzyme's active site. The human serine/threonine
XX protein kinase and its fragments are also useful in raising specific
XX antibodies that can block the enzyme and effectively reduce its activity.
XX Human serine/threonine protein kinase sequences can be used in the
XX preparation of a medicament for modulating the activity of a serine/
XX threonine protein kinase in a disease, e.g. cancer, diabetes, a central
XX nervous system (CNS) disorder, a respiratory disorder (including chronic
XX obstructive pulmonary disease), a cardiovascular disorder, a
XX dermatological disorder, a gastrointestinal or liver disease, a
XX haematological disorder, a musculoskeletal disorder, a reproductive
XX disorder, or a urological disorder. CNS disorders may include Alzheimer's
XX disease, Parkinson's disease, stroke or neuropathic pain
XX
XX
XX Sequence 1584 BP; 418 A; 376 C; 387 G; 403 T; 0 U; 0 Other;
Query Match 100.0%; Score 1308; DB 9; Length 1584;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

241 CAGGAAGAGATCTGAGGAGGAGGTTTCACTGAGAGAGAGCTCCCTTTGGGAGGCC 300
301 TCATCTTACCTTGAAGTGGAGAGCTGAGTGAAGCTCTTATGCGACAGTTTACAAGGG 360
301 TCATCTTACCTTGAAGTGGAGAGCTGAGTGAAGCTCTTATGCGACAGTTTACAAGGG 360
361 ATTAGAGAAATTAATGAGCACTAGTGGCTTTAAAAGTATCAGCATGAAATGAGAGAA 420
361 ATTAGAGAAATTAATGAGCACTAGTGGCTTTAAAAGTATCAGCATGAAATGAGAGAA 420
421 GGAGTCCCATTTACAGCTATCCGAGAGGCTTCTCCCTGGAAGGTTTGAAGATGCAAT 480
421 GGAGTCCCATTTACAGCTATCCGAGAGGCTTCTCCCTGGAAGGTTTGAAGATGCAAT 480
481 ATTGTCCTCTGACATGACATAATTCACACCAAGAGACATGACATGTTTGAATAC 540
481 ATTGTCCTCTGACATGACATAATTCACACCAAGAGACATGACATGTTTGAATAC 540
541 ATGACACAGACCTGGCCGATATATGCTCAGCATCCAGAGAGGCTTATCTCATTAAT 600
541 ATGACACAGACCTGGCCGATATATGCTCAGCATCCAGAGAGGCTTATCTCATTAAT 600
601 GTCAAGCTTTGATGTTCACTTTTGGGAGGCTGAGGCTGATCATGACACCAACACGTT 660
601 GTCAAGCTTTGATGTTCACTTTTGGGAGGCTGAGGCTGATCATGACACCAACACGTT 660
661 CTTCACAGGACCTGAAACCTGAGAACTTATCATGATCAGTCACTGAGAGGCTCAACTG 720
661 CTTCACAGGACCTGAAACCTGAGAACTTATCATGATCAGTCACTGAGAGGCTCAACTG 720
721 GGTGATTTTGGCTTGGCCCGGCGCAAGTCCATCCACAGACATCACTCTTGAAGAGC 780
721 GGTGATTTTGGCTTGGCCCGGCGCAAGTCCATCCACAGACATCACTCTTGAAGAGC 780
781 GTGACCTCTGATGATCGGCGCCCTGATGCTTGTGAGAGCAGTGAATATCTCTGAG 840
781 GTGACCTCTGATGATCGGCGCCCTGATGCTTGTGAGAGCAGTGAATATCTCTGAG 840
841 CTGACATATGAGGGGTGACAGGCTGACATCTTATGAAATGTTCCAGGCTCAACCTTGT 900
841 CTGACATATGAGGGGTGACAGGCTGACATCTTATGAAATGTTCCAGGCTCAACCTTGT 900
901 CCGGGGTTTCCAAATCCTTGAACAGCTGAGAGAAATCTGGAGAGTCTGGAGATCCCT 960
901 CCGGGGTTTCCAAATCCTTGAACAGCTGAGAGAAATCTGGAGAGTCTGGAGATCCCT 960
961 ACAGAGATATCTTGGCGGAGGTCTCAAGCTTAACTTAATCAATCCAGAAATGTTCCCA 1020
961 ACAGAGATATCTTGGCGGAGGTCTCAAGCTTAACTTAACTTAATCAATCCAGAAATG 1020
1021 CTGCTAGGCTCGAAGCTTATGTTGTTGGAACAGGCTGAGGCTGAGGTTCTGAACT 1080
1021 CTGCTAGGCTCGAAGCTTATGTTGTTGGAACAGGCTGAGGCTGAGGTTCTGAACT 1080
1081 GAAAGACCTGGCTCCCAAGATCTAAAGGCTTTCCCAAGACCGGCTCTCGCCCAAGAA 1140
1081 GAAAGACCTGGCTCCCAAGATCTAAAGGCTTTCCCAAGACCGGCTCTCGCCCAAGAA 1140
1141 GCACTTGTTCATGATTAATTTAGAGCCCTGCGCATCTCAGTGTACCAAGCTTCTGATAG 1200
1141 GCACTTGTTCATGATTAATTTAGAGCCCTGCGCATCTCAGTGTACCAAGCTTCTGATAG 1200
1201 GAGTCTTTGTTTACAGTTTCAAGAGTGAAGCTTAAAGCCAGAAATGTTGAGACTTTGGCC 1260
1201 GAGTCTTTGTTTACAGTTTCAAGAGTGAAGCTTAAAGCCAGAAATGTTGAGACTTTGGCC 1260
1261 TCCTACAGAAAGATCACACCCAGCCAGTTTATGCAATATCTGTGA 1308
1261 TCCTACAGAAAGATCACACCCAGCCAGTTTATGCAATATCTGTGA 1308
RESULT 4
ACC79971

XX	ID	ACC79971 standard; cDNA; 1628 BP.
XX	AC	ACC79971;
XX	DT	09-SEP-2003 (first entry)
DE		Human serine/threonine protein kinase encoding cDNA SEQ ID NO:9.
KM		Human; serine/threonine protein kinase; kinase; enzyme; cytosolic; antidiabetic; neuroprotective; nootropic; antiparkinsonian; analgesic; KMEuroprotective; antiinflammatory; gastroenterostinal; hepatotrophic; XX gynaecological; uropathic; dermatological; gene therapy; cancer; KM diabetes; central nervous system disorder; CNS disorder; liver disease; respiratory disorder; chronic obstructive pulmonary disease; stroke; KM cardiovascular disorder; dermatological disorder; urological disorder; gastroenterostinal disease; haematological disorder; Alzheimer's disease; KM musculoskeletal disorder; reproductive disorder; Parkinson's disease; neuropathic pain; gene; ss.
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FT	CDS	45..1352
FT	/ftaga=	a "serine/threonine protein kinase"
FT	/product=	"serine/threonine protein kinase"
XX	MO2003046167-A1.	
PD	05-JUN-2003.	
XX	26-NOV-2002; 2002WO-BP033260.	
Pf	27-NOV-2001; 2001US-0333131P.	
XX	(FARB) BAYER AG.	
PA	Koehler RH;	
PI	WPI; 2003-505196/47.	
DR	P-PADB; ABR57361.	
XX	New polynucleotide encoding a serine/threonine protein kinase PT polypeptide, useful for diagnosing, preventing or treating diseases PT associated with serine/threonine protein kinase dysfunction, e.g. cancer PT or diabetes.	
PS	Claim 1; Page 185-188; 196pp; English.	
CC	The present sequence encodes a human serine/threonine protein kinase (I). CC (1) has cytostatic, antidiabetic, neuroprotective, antiparkinsonian, CC nootropic, cerebroprotective, analgesic, antiinflammatory, hepatotrophic, CC gastroenterostinal, gynaecological, uropathic and dermatological activities, and can be used in gene therapy. Serine/threonine protein CC kinase polynucleotide and polypeptide sequences can be used in diagnosing, preventing, ameliorating or treating diseases associated with CC serine/threonine protein kinase dysfunction. They may also be used to identify test compounds that may act, for example, as activators or inhibitors at the enzyme's active site. The human serine/threonine protein kinase and its fragments are also useful in raising specific CC antibodies that can block the enzyme and effectively reduce its activity. CC Human serine/threonine protein kinase sequences can be used in the CC preparation of a medicament for modulating the activity of a serine/ threonine protein kinase in a disease, e.g. cancer, diabetes, a central nervous system (CNS) disorder, a respiratory disorder (including chronic obstructive pulmonary disease), a cardiovascular disorder, a dermatological disorder, a gastroenterostinal or liver disease, a haematological disorder, a musculoskeletal disorder, a reproductive disorder, or a urological disorder. CNS disorders may include Alzheimer's disease, Parkinson's disease, stroke or neuropathic pain	
SO	Sequence 1628 BP; 429 A; 383 C; 400 G; 416 T; 0 U; 0 Other;	
Query Match	100.0%; Score 1308; DB 9; Length 1628;	

	Best Local Matches 1308	Similarity Conservative	100.0%	Pred. No. 0	Mismatches 0	Indels 0	Gaps 0
QY	1	ATGGGTGTCAGAGCTGTGTGTCAGAAAGCTGTACAGGCTTGGATGTCAGCTGTCTACCATTTGTTCA	60				
Db	45	ATGGGTCAAGAGCTGTGTGTCAGAAAGCTGTACAGCTTGGATGTCAGCTGTCTACCATTTGTTCA	104				
QY	61	GAGGAGGCGGACGACACAGCTGTCTGGAGAGTACGCTTGAGACACGAGAGCTTCCGCTTC	120				
Db	105	GAGGAGGCGGACGACACAGCTGTCTGGAGAGTACGCTTGAGACACGAGAGCTTCCGCTTC	164				
QY	121	AAGCTACAGACCCATAAAGAGCATCTCATGTTCCATGACTTCATTTCACTCCGAGGGGACTT	180				
Db	165	AAGCTAACAGACCTTAAGAAGCATCATGTTCCATGACTTCATTTCACTCCGAGGGGACTT	224				
QY	181	CAAGCTCCCGCTGCCACAGAACTTCAAGAGTAAAGGCGCACGAGATTAACAGTATTTGTTT	240				
Db	225	CAAGCTCCCGCTGCCACAGAACTTCAAGAGTAAAGGCGCACGAGATTAACAGTATTTGTTT	284				
QY	241	CAGGAAGAGATCTGAGGCAAGGCTTTTCAGTGGAGAAAGGCTTCCTTTTGGGCGACCC	300				
Db	285	CAGGAAGAGATCTGAGGCAAGGCTTTTCAGTGGAGAAAGGCTTCCTTTTGGGCGACCC	344				
QY	301	TCACTTAACTTTGAACCTTGGAGAGCTGGGTGGAAGGCTCTTAATGGACAGCTTTTACAAGGGG	360				
Db	345	TCACTTAACTTTGAACCTTGGAGAGCTGGGTGGAAGGCTCTTAATGGACAGCTTTTACAAGGGG	404				
QY	361	ATTAGCAGAAATTAATATGACAACTAGTGGCTTTAAAGATCAATCAGCATGAAATGACAGAGAA	420				
Db	405	ATTAGCAGAAATTAATATGACAACTAGTGGCTTTAAAGATCAATCAGCATGAAATGACAGAGAA	464				
QY	421	GGAATCCCATTTTACAGGTATCCGAGAGCTTCTCTCTGAAAGGCTTTGAAACATGCCAAT	480				
Db	465	GGAATCCCATTTTACAGGTATCCGAGAGCTTCTCTCTGAAAGGCTTTGAAACATGCCAAT	524				
QY	481	ATTGTGCTCTGTCATGACATTAATCCACACCAAGAGACCTGACATTCGTTTTTGAATAC	540				
Db	525	ATTGTGCTCTGTCATGACATTAATCCACACCAAGAGACCTGACATTCGTTTTTGAATAC	584				
QY	541	ATGCACACAGACCTGGGCGCCAGTATATGCTCTCAGGATCAAGGAGGCTTCACTCTCATTAAT	600				
Db	585	ATGCACACAGACCTGGGCGCCAGTATATGCTCTCAGGATCAAGGAGGCTTCACTCTCATTAAT	644				
QY	601	GTCAGACTTTTCATGTTTCAACTTTTGGGCGGCTTGGCGTACATCCACACCAACACGTT	660				
Db	645	GTCAGACTTTTCATGTTTCAACTTTTGGGCGGCTTGGCGTACATCCACACCAACACGTT	704				
QY	661	CTTTCACAGGGAACCTGAAACCTCAGAACTTACTCATCTGATCACCTGGGAGAGCTCAAACTG	720				
Db	705	CTTTCACAGGGAACCTGAAACCTCAGAACTTACTCATCTGATCACCTGGGAGAGCTCAAACTG	764				
QY	721	GCTGATTTTGGTCTGTCGCGGGGCCAATCCATTCGCCAGCCAGACATACCTCTTCAGAAAGTC	780				
Db	765	GCTGATTTTGGTCTGTCGCGGGGCCAATCCATTCGCCAGCCAGACATACCTCTTCAGAAAGTC	824				
QY	781	GTCACCTCTGTGTACCGGCGCCCTGTGATGCTTTTGTCTGGAGGCACTGTAAATTTCTCTGAG	840				
Db	825	GTCACCTCTGTGTACCGGCGCCCTGTGATGCTTTTGTCTGGAGGCACTGTAAATTTCTCTGAG	884				
QY	841	CTGGACATATGGGGGTGACGGCTGTACTTTATTTGAAATGTTCCAGGGTCAACTTTGTTT	900				
Db	885	CTGGACATATGGGGGTGACGGCTGTACTTTATTTGAAATGTTCCAGGGTCAACTTTGTTT	944				
QY	901	CTGGGGGTTTCCAAACATCTTTGAAACAGCTGTGAGAAATCTGGGAGGGGTCTGGGAGTCCCT	960				
Db	945	CTGGGGGTTTCCAAACATCTTTGAAACAGCTGTGAGAAATCTGGGAGGGGTCTGGGAGTCCCT	1004				
QY	961	ACAGAGGATACCTTGGCGGGAGTCTCCAACTACCTTAACATCCAGATGAGTTTCCCA	1020				
Db	1005	ACAGAGGATACCTTGGCGGGAGTCTCCAACTACCTTAACATCCAGATGAGTTTCCCA	1064				
QY	1021	CTGGCTTACGCTCTGAAAGCTTTCATGTTGTCTGGAACAGGCTGGGCAAGGCTTCTGAAAGCT	1080				

Db 1065 CTGCTACGCTCGAAGCTTCATGTTGTCGGAACAGGCTGGAGGTTCTGAAGCT 1124
QY 1081 GAAAGCTGCGCTCCAGATGCTAAAAGGCTTTTCCAGAGACCGCTCTCCGCCAGAA 1140
Db 1125 GAAGACCTGCGCTCCAGATGCTAAAAGGCTTTTCCAGAGACCGCTCTCCGCCAGAA 1184
QY 1141 GCACTGCTCATATATTTTTCAGGCGCCCTCCCATCTGAGCTGTAACAGCTTCTGATGAG 1200
Db 1185 GCACTGCTCATATATATTTTTCAGGCGCCCTCCCATCTGAGCTGTAACAGCTTCTGATGAG 1244
QY 1201 GAGCTCTTGTTCACAGTTTTCAGAGAGTGAAGCTTAAAGCCAGAAATGTTGACCTTTGGCC 1260
Db 1245 GAGCTCTTGTTCACAGTTTTCAGAGAGTGAAGCTTAAAGCCAGAAATGTTGACCTTTGGCC 1304
QY 1261 TCCTACCAAGAAAGTCAACCAAGCCCAAGTTTACCAATGCTGTGA 1308
Db 1305 TCCTACCAAGAAAGTCAACCAAGCCCAAGTTTACCAATGCTGTGA 1352

RESULT 5

ADJ96563
ID ADJ96563 standard; DNA; 2250 BP.

AC ADJ96563;

DT 06-MAY-2004 (first entry)

XX Human cyclin dependent kinase PFTAIR2 DNA SeqID 20.

XX Gene; de; kinase; human; tyrosine protein kinase;

KW serine/threonine protein kinase; PTK; STK; gene therapy; cancer;

KW immune-related disease; cardiovascular disease; brain;

KW neuronal associated disease; metabolic; inflammatory disorder;

KW cytosolic; neuroprotective; immunomodulator; antiinflammatory;

XX cyclin dependent kinase; PFTAIR2.

XX Homo sapiens.

OS 51.

XX MO2004006838-A2.

PN 22-JAN-2004.

PD 15-JUL-2003; 2003WO-US021730.

PF 15-JUL-2002; 2002US-035632P.

PI (SUGEN-) SUGEN INC.

XX Whyte D, Manning G, Caenepeel S;

XX MPI; 2004-122753/12.

XX P-PSDB; ADJ965629.

XX New nucleic acid molecule encoding a kinase polypeptide, useful for

XX PT preparing a composition for treating diseases or disorders, e.g., cancer,

XX PT or neurological, immunological or inflammatory disorders.

XX Example 1; SEQ ID NO 20; 366bp; English.

XX This invention relates to a novel isolated, enriched or purified nucleic

XX acid molecule that encodes a kinase polypeptide. Specifically, it relates

XX to human tyrosine and serine/threonine protein kinases (PTK's and STK's),

XX as well as protein kinase-like enzymes. The present invention describes

XX screening methods to identify agonists, antagonists and antibodies that

XX can be used to modulate the activity or function of the mammalian kinase

XX enzymes. As such, these compositions can be used for gene therapy

XX purposes to treat diseases or disorders including cancer, immune-related

XX diseases, cardiovascular disease, brain or neuronal associated disease,

XX metabolic and inflammatory disorders. Accordingly, they exhibit

XX cytosolic, neuroprotective, immunomodulator and antiinflammatory

XX activities. This polynucleotide sequence is a human kinase DNA sequence

XX of the invention.

XX Sequence 2250 BP; 616 A; 494 C; 531 G; 609 T; 0 U; 0 Other;

SQ Query Match 100.0%; Score 1308; DB 12; Length 2250;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTCAAGAGCTGTGTCAGAAAGCTGACAGCTGTGATGAGCTGCTACATTTGCA 60
Db 45 ATGGGTCAAGAGCTGTGTCAGAAAGCTGACAGCTGTGATGAGCTGCTACATTTGCA 104
QY 61 GAGGAGGCGGAGGACACACAGCTGTGAGGAGTCAAGCTGAGACCAAGAGCTGCTTC 120
Db 105 GAGGAGGCGGAGGACACACAGCTGTGAGGAGTCAAGCTGAGACCAAGAGCTGCTTC 164
QY 121 AAGCTACAGAGCTTAAAGAGACATCATGTTTCATGACTTTCATTTCAAGGAGGACT 180
Db 165 AAGCTACAGAGCTTAAAGAGACATCATGTTTCATGACTTTCATTTCAAGGAGGACT 224
QY 181 CAAGTGCCTGTCGCCAGAAAGTTCAAGAGTAAAGGCCACGAGTAAACAGTATTTT 240
Db 225 CAAGTGCCTGTCGCCAGAAAGTTCAAGAGTAAAGGCCACGAGTAAACAGTATTTT 284
QY 241 CAGGAAAGAGATCTGAGCAGGCTTTTCAGTGAAGAAAGGCTCTTTGGGACAGCC 300
Db 285 CAGGAAAGAGATCTGAGCAGGCTTTTCAGTGAAGAAAGGCTCTTTGGGACAGCC 344
QY 301 TCATCTTACTTGAAGCTTGAAGAGCTGAGGCTCTTATGACAGTTTAAAGGGG 360
Db 345 TCATCTTACTTGAAGCTTGAAGAGCTGAGGCTCTTATGACAGTTTAAAGGGG 404
QY 361 ATTACAGATTAATGAGCACTAGTGTGTTTAAAGTATCAGATGAATGACAGGAA 420
Db 405 ATTACAGATTAATGAGCACTAGTGTGTTTAAAGTATCAGATGAATGACAGGAA 464
QY 421 GGAGTCCCATTTACAGACTATCCGAGAGGCTTCTCTGAAGGGTTTGAACAATGCAAT 480
Db 465 GGAGTCCCATTTACAGACTATCCGAGAGGCTTCTCTGAAGGGTTTGAACAATGCAAT 524
QY 481 ATTGTGCTCTGATGATGATATTCACACCAAGAGACATGACATTTGTTGAAATAC 540
Db 525 ATTGTGCTCTGATGATGATATTCACACCAAGAGACATGACATTTGTTGAAATAC 584
QY 541 ATGCAACAGAGCTGAGCCAGTATATGTCAGACATCCAGAGGCTTCATCTATAT 600
Db 585 ATGCAACAGAGCTGAGCCAGTATATGTCAGACATCCAGAGGCTTCATCTATAT 644
QY 601 GTGACATTTTCATGATTTCACTTTTGGGGGCTGGGCTACATCCACCAACAGCTT 660
Db 645 GTGACATTTTCATGATTTCACTTTTGGGGGCTGGGCTACATCCACCAACAGCTT 704
QY 661 CTTCACAGGAGCTGAAACCTCAGAACTTACTCATGATCACTGGAGAGCTCAAACTG 720
Db 705 CTTCACAGGAGCTGAAACCTCAGAACTTACTCATGATCACTGGAGAGCTCAAACTG 764
QY 721 GGTATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 765 GGTATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 824
QY 781 GTGACCTCTGTGATACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 825 GTGACCTCTGTGATACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 884
QY 841 CTGACATATGAGGCTGAGCTGATCTTATATGAAATGTTCCAGGGCTCAACTTTGTTT 900
Db 885 CTGACATATGAGGCTGAGCTGATCTTATATGAAATGTTCCAGGGCTCAACTTTGTTT 944
QY 901 CCTGGGGTTTCCACATCTTGAACAGCTGAGGAAATCTGGAGGCTGCTGGAGTCCCT 960
Db 945 CCTGGGGTTTCCACATCTTGAACAGCTGAGGAAATCTGGAGGCTGCTGGAGTCCCT 1004
QY 961 ACAGAGATTAATGAGCTGAGGAGTCTCAAGCTTACTAATCAATCCAGAAATGTTCCCA 1020

DB 1005 ACAGAGATACCTTGGCCGGAGTCTCCAGCTACCTACATCAATCCAGAAATGTTCCCA 1064
 QY 1021 CTGCTACGCTCTGGAAGCTTCATGTTGCTGGAACAGGCTGGGCAAGGTTCTGAAGCT 1080
 DB 1065 CTGCTACGCTCTGGAAGCTTCATGTTGCTGGAACAGGCTGGGCAAGGTTCTGAAGCT 1124
 QY 1081 GAAGACCTGGCCCTCCAGATGCTAAAGAGCTTCCAGAGACCGGCTCCCGCCAGAGAA 1140
 DB 1125 GAAGACCTGGCCCTCCAGATGCTAAAGAGCTTCCAGAGACCGGCTCCCGCCAGAGAA 1184
 QY 1141 GCACTGTTCAATATTTTCAAGGCTGCTGCACTCTCAGCTGTACAGCTTCTGATGAG 1200
 DB 1185 GCACTGTTCAATATTTTCAAGGCTGCTGCACTCTCAGCTGTACAGCTTCTGATGAG 1244
 QY 1201 GACTCTTTGTTTACAGTTTCAGAGAGTGAAGCTTAAAGCCAGAAATGTGTGACCTTTTGCC 1260
 DB 1245 GAGTCTTTGTTTACAGTTTCAGAGAGTGAAGCTTAAAGCCAGAAATGTGTGACCTTTTGCC 1304
 QY 1261 TCCTACAGAAAGTGCACACCCAGCCAGCTTATAGCAAAATGCTGTGA 1308
 DB 1305 TCCTACAGAAAGTGCACACCCAGCCAGCTTATAGCAAAATGCTGTGA 1352

RESULT 6

ADL22554
 ID ADL22554 standard; cDNA, 2140 BP.

XX ADL22554;
 AC
 XX

XX 20-MAY-2004 (first entry)
 DT
 XX

DE Human disease detection and treatment (MDPT) cDNA - SEQ ID 3.
 XX

KW disease detection; MDP; antiarteriosclerotic; antipsoriatic; cytoskeletal;
 cell signalling; arteriosclerosis; psoriasis; cancer; human; ss; gene.
 XX

OS Homo sapiens.
 XX

XX WO2003062379-A2.
 PN

XX 31-JUL-2003.
 PD

XX 14-JAN-2003; 2003MO-US001363.
 PF

XX 17-JAN-2002; 2002US-0349413P.
 PR

XX 17-JAN-2002; 2002US-0349946P.
 PR

XX (INCY-) INCYTE GENOMICS INC.
 PA

PI Jones AL, Dahl CR, Gietzen D, Chinn J, Dufour GE, Jackson JL,
 PI Yu JY, Tusson O, Yap PE, Ameshey SR, Dam TC, Liu TF, Gerstein EH,
 PI Peralta CH, Lewis SA, Chen AJ, Marwaha R, Lan RX, Urashka MB,
 PI Kristnam SR, Kolluru V, Panesar IS;
 PI

DR MPI, 2003-853443/79.
 DR

DR P-PSDB; ADL22658.
 DR

PT New isolated disease detection and treatment polynucleotide for
 PT diagnosing or treating conditions associated with cell signaling e.g.
 PT arteriosclerosis, psoriasis, and cancer.
 PT

XX Claim 1; SEQ ID NO 3; 411bp; English.
 PS

XX The invention relates to a novel isolated disease detection and treatment
 CC (MDPT) DNA polynucleotide. The polynucleotide of the invention
 CC demonstrates antiarteriosclerotic, antipsoriatic and cytoskeletal
 CC activities and may be useful in a composition for detecting the
 CC expression of a disease detection and treatment molecule polynucleotide.
 CC The molecules of the invention may be utilised to diagnose or treat
 CC conditions, diseases or disorders associated with cell signalling, such
 CC as arteriosclerosis, psoriasis and cancer. The current sequence is that
 CC of a human MDP cDNA of the invention.
 XX

SO Sequence 2140 BP; 601 A; 485 C; 493 G; 561 T; 0 U; 0 Other;

Query Match 91.1%; Score 1191; DB 11; Length 2140;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1239; Conservative 0; Mismatches 0; Indels 8; Gaps 4;

QY 70 GAGGACACAGGCTGTGGAGAGTCAAGCTGAGACACAGGAGGCTGCTCAAGCTAAACA 129
 DB 1 GAGGACACAGGCTGTGGAGAGTCAAGCTGAGACACAGGAGGCTGCTCAAGCTAAACA 60
 QY 130 GACCTTAAAGAGCATCATGTTTCAATGACTTATTTTCAAGGAGGAGCTTCAAGCTGCC 189
 DB 61 GACCTTAAAGAGCATCATGTTTCAATGACTTATTTTCAAGGAGGAGCTTCAAGCTGCC 120
 QY 190 CGTGCCAGAGGTTCAAGAGTAAAGGCGACGAGTAAAGTATGTTTTCAGAGAGAG 249
 DB 121 CGTGCCAGAGGTTCAAGAGTAAAGGCGACGAGTAAAGTATGTTTTCAGAGAGAG 180
 QY 250 GATCTGAGGCGAGGTTTTCAGAG-----GAGGAGAGGCTCTTTTGGGGAGGCTCAT 304
 DB 181 GATCTGAGGCGAGGTTTTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
 QY 305 CTTACTTGAAGCTTGAAGAGTGGGAGAGGCTCTTATGCAAGATTTCAGAGGAGATTA 364
 DB 241 CTTACTTGAAGCTTGAAGAGTGGGAGAGGCTCTTATGCAAGATTTCAGAGGAGATTA 300
 QY 365 GCAGAAATTAATGACAACTAGTGGCTTTTAAAGTCAAGATGAATGAGAGAGAGAG 424
 DB 301 GCAGAAATTAATGACAACTAGTGGCTTTTAAAGTCAAGATGAATGAGAGAGAGAGAG 360
 QY 425 TCCCATTTACAGTATCCGAGAGGCTTCTCTCTGAGAGGTTTGAACATGCCCAATATG 484
 DB 361 TCCCATTTACAGTATCCGAGAGGCTTCTCTCTGAGAGGTTTGAACATGCCCAATATG 420
 QY 485 TGCTCTGATGACATTAATCCACACCAAGAGACATGCTGTTTGAATACATGC 544
 DB 421 TGCTCTGATGACATTAATCCACACCAAGAGACATGCTGTTTGAATACATGC 480
 QY 545 ACACAGACCTGGCCCGATATATGCTCAGCATCCAGAGAGGCTTCATCTCATTAATGCA 604
 DB 481 ACACAGACCTGGCCCGATATATGCTCAGCATCCAGAGAGGCTTCATCTCATTAATGCA 540
 QY 605 GACTTTTCAATGTTTCACTTTTGGGGGGCTGGCGGTCAATCCACACCC-AAACGTTCTT 663
 DB 541 GACTTTTCAATGTTTCACTTTTGGGGGGCTGGCGGTCAATCCACACCC-AAACGTTCTT 600
 QY 664 CACAGGAGCTGTAAGAACTCAGAACTTACTCATGATCACTGGAGAGGCTCAAACTGGCT 723
 DB 601 CACAGGAGCTGTAAGAACTCAGAACTTACTCATGATCACTGGAGAGGCTCAAACTGGCT 660
 QY 724 GATTT-TGGTCTTGGCCCGGCAAGTCCATTCCAGCCAGACATCTTTCAGAGTGT 782
 DB 661 GATTTGTGTTCTTGGCCCGGCAAGTCCATTCCAGCCAGACATCTTTCAGAGTGT 720
 QY 783 GACCTCTGTGACCGGCCCCCTGATGCTTGTGGAGGACCATGAATATCTCTGAGCT 842
 DB 721 GACCTCTGTGACCGGCCCCCTGATGCTTGTGGAGGACCATGAATATCTCTGAGCT 780
 QY 843 GGACATATAGGGGTGAGAGCTGATCTTTATTAAGATTTCCAGGGGCAACTTGTGTC 902
 DB 781 GGACATATAGGGGTGAGAGCTGATCTTTATTAAGATTTCCAGGGGCAACTTGTGTC 840
 QY 903 TGGGGTTTCCAAATCTTGAACAGCTGAGAGAAATCTGGAGGTGCTGGAGTCCCTTAC 962
 DB 841 TGGGGTTTCCAAATCTTGAACAGCTGAGAGAAATCTGGAGGTGCTGGAGTCCCTTAC 900
 QY 963 AAGAGATATTTGGCCGGAGGTCTCAAGCTACTTCAATCAATCCAGAAATGTTCCACT 1022
 DB 901 AAGAGATATTTGGCCGGAGGTCTCAAGCTACTTCAATCAATCCAGAAATGTTCCACT 960
 QY 1023 GCCTACGCTCGAAGGCTTCAATGTTGCTGGAACAGGCTGGGAGGCTTCTGAAGCTGA 1082
 DB 961 GCCTACGCTCGAAGGCTTCAATGTTGCTGGAACAGGCTGGGAGGCTTCTGAAGCTGA 1020

QY 1083 AGACCTGGCTCCAGATGTATAAGCTTTCCAGAGACCGGCTCTCCGCCAGGAAGC 1142
DB 1021 AGACCTGGCTCCAGATGTATAAGCTTTCCAGAGACCGGCTCTCCGCCAGGAAGC 1080
QY 1143 ACTTGTCATGATATATTTGAGCGCCCTGCATCTCAGCTGTACAG-CTTCTGTATGAG 1201
DB 1081 ACTTGTCATGATATATTTGAGCGCCCTGCATCTCAGCTGTACAGACTTCCGTGATGAG 1140
QY 1202 AGCTTTGTTTACAGTTTCAGAGGTAGAGCTTAAGCCAGAAATGTGTACCTTTGACCT 1261
DB 1141 AGCTTTGTTTACAGTTTCAGAGGTAGAGCTTAAGCCAGAAATGTGTACCTTTGACCT 1200
QY 1262 CCACACGAAAGGTACACCCAGCCAGTTTACGAAATGCTGTGA 1308
DB 1201 CCTACGAAAGGTACACCCAGCCAGTTTACGAAATGCTGTGA 1247

RESULT 7
ADCC0755
ID ADCC0755 standard; cDNA; 1311 BP.
XX
AC ADCC0755;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel cDNA sequence, SEQ ID NO:837.
XX
XX Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; anti-nausea; anticoagulant; thrombolytic; vulnerary;
KW anticancer; osteopathic; immunosuppressive; anti-inflammatory; cytostatic;
KW gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002MO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
XX MPI; 2003-371981/35.
DR P-SDB; ADCC1726.
XX
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Claim 1; SEQ ID NO 837; 1185bp; English.
XX
XX The invention relates to 971 novel human cDNA sequences (ADCC29919-
CC ADCC0889) and the polypeptides they encode (ADCC0890-ADCC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or

CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADCC1861-ADCC32627) and the polypeptides encoded by the contigs (ADCC32628
CC -ADCC3394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 1311 BP; 332 A; 328 C; 329 G; 322 T; 0 U; 0 Other;

Query Match 90.7%; Score 1186; DB 10; Length 1311;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;

Matches 1186; Conservative 0; Indels 0;

QY 123 GCTAACAGACCTTAAGAAAGCATCATGTTCCATGATCTTCAATTCACCCAGGGGACTTCA 182
DB 72 GCTAACAGACCTTAAGAAAGCATCATGTTCCATGATCTTCAATTCACCCAGGGGACTTCA 131
QY 183 AGCTGCCCTGCCAGAACTTCAAGATTAAGGCGCACCGAATTAACGATTTGTTTCA 242
DB 132 AGCTGCCCTGCCAGAACTTCAAGATTAAGGCGCACCGAATTAACGATTTGTTTCA 191
QY 243 GGAAGAGATCTGAGGCGGGTTTCAGTGGAGGAAGCCCTCTTTGGGGAGCCTC 302
DB 192 GGAAGAGATCTGAGGCGGGTTTCAGTGGAGGAAGCCCTCTTTGGGGAGCCTC 251
QY 303 ATCTTACTTGAATCTGAGAAAGCTGGGTGAAGGCTTTATGCGACATTTTACAGGGGAT 362
DB 252 ATCTTACTTGAATCTGAGAAAGCTGGGTGAAGGCTTTATGCGACATTTTACAGGGGAT 311
QY 363 TAGCAGATTAATGGAACAATGAGCTTTTAAAGTCATCAGATGAATGCAAGGAAG 422
DB 312 TAGCAGATTAATGGAACAATGAGCTTTTAAAGTCATCAGATGAATGCAAGGAAG 371
QY 423 AGTCCATTATTAAGGATCCGAGAAAGCTCTCCGGAAGGGTTGGAACATGCCAATAT 482
DB 372 AGTCCATTATTAAGGATCCGAGAAAGCTCTCTCGGAAGGGTTGGAACATGCCAATAT 431
QY 483 TGTGCTCTGATGATCATATATCCACCAAGAGACATGATTCGTTTGAATATAT 542
DB 432 TGTGCTCTGATGATCATATATCCACCAAGAGACATGATTCGTTTGAATATAT 491
QY 543 GCAACAGACTGGGCCAGTATATGTTCAAGATCCAGAGGGCTTCATCTCATATAT 602
DB 492 GCAACAGACTGGGCCAGTATATGTTCAAGATCCAGAGGGCTTCATCTCATATAT 551
QY 603 CAGACTTTTCATGTTTCAACTTTTGGGGGCTGGGGTATCCACCAACAGCTTCT 662
DB 552 CAGACTTTTCATGTTTCAACTTTTGGGGGCTGGGGTATCCACCAACAGCTTCT 611
QY 663 TCACAGGAGCTGAACCTCGAAGCTTACTCATCATGCTCGGAGAGACTCAAACTGCG 722
DB 612 TCACAGGAGCTGAACCTCGAAGCTTACTCATCATGCTCGGAGAGACTCAAACTGCG 671
QY 723 TGAATTTGCTCTGCGCGGCGCAAGTCATTTCCAGCGACGACATATCTTTCAGAACTGCT 782
DB 672 TGAATTTGCTCTGCGCGGCGCAAGTCATTTCCAGCGACGACATATCTTTCAGAACTGCT 731

```

OY 783 GACCTCTGTGTACCGGCCCCCTGATGCTTGTCTGGAGACCACTGAATATTCTCTGAGCT 842
DB 732 GACCCCTGTGTACCGGCCCCCTGATGCTTGTCTGGAGACCACTGAATATTCTCTGAGCT 791
OY 843 GGACATATGAGGTGACGAGGTGATCTTTATTTGAAATGTTCCAGGGTCACTTTGTTTCC 902
DB 792 GACATATGAGGTGACGAGGTGATCTTTATTTGAAATGTTCCAGGGTCACTTTGTTTCC 851
OY 903 TGGGGTTTCCAACTCTTTGAAACAGCTGAGAAATCTGGAGAGTCTGGAGATCCCTAC 962
DB 852 TGGGGTTTCCAACTCTTTGAAACAGCTGAGAAATCTGGAGAGTCTGGAGATCCCTAC 911
OY 963 AGAGATATCTTGGCCGGGAGTCTCCAGACTACTTCAATCCAGATGTTTCCCACT 1022
DB 912 AGAGATATCTTGGCCGGGAGTCTCCAGACTACTTCAATCCAGATGTTTCCCACT 971
OY 1023 GCTTACGCTCGAGGCTTCTGTTGTCTGAAACAGGCTGGGAGGGTCTGAAAGCTGA 1082
DB 972 GCTTACGCTCGAGGCTTCTGTTGTCTGAAACAGGCTGGGAGGGTCTGAAAGCTGA 1031
OY 1083 AGACCTGGCTTCCAGATGCTTAAAGGCTTTCCAGAGACCGGCTCTCCGCGCAGGAGAC 1142
DB 1032 AGACCTGGCTTCCAGATGCTTAAAGGCTTTCCAGAGACCGGCTCTCCGCGCAGGAGAC 1091
OY 1143 ACTTGTTCATGATTTATTTAGAGGCTCTGCACTCTGAGCTGTACCACTTCTGATAGAGA 1202
DB 1092 ACTTGTTCATGATTTATTTAGAGGCTCTGCACTCTGAGCTGTACCACTTCTGATAGAGA 1151
OY 1203 GCTTTGTTTACAGTTTTCAGAGTGTAGAGGCTTAAAGCAGAAATGTGTGACCTTTGGGCTTC 1262
DB 1152 GCTTTGTTTACAGTTTTCAGAGTGTAGAGGCTTAAAGCAGAAATGTGTGACCTTTGGGCTTC 1211
OY 1263 CTACAGAAAGGTCAACCAACGAGCCCAAGTTTAGCAAAATGCTGTGTA 1308
DB 1212 CTACAGAAAGGTCAACCAACGAGCCCAAGTTTAGCAAAATGCTGTGTA 1257

```

RESULT 8

ACCT79969

ID ACCT79969 standard; cDNA, 1534 BP.

ACCT79969;

DT 09-SEP-2003 (first entry)

DE Human serine/threonine protein kinase nucleotide sequence SEQ ID NO:5.

Human, serine/threonine protein kinase; kinase; enzyme; cytosolic;
 antidiabetic; neuroprotective; nootropic; antiparkinsonian; analgesic;
 cerebroprotective; antiinflammatory; gastrointestinal; hepatotropic;
 gynaecological; uropathic; dermatological; gene therapy; cancer;
 diabetes; central nervous system disorder; CNS disorder; liver disease;
 respiratory disorder; chronic obstructive pulmonary disease; stroke;
 cardiovascular disorder; dermatological disorder; urological disorder;
 gastrointestinal disease; haematological disorder; Alzheimer's disease;
 musculoskeletal disorder; reproductive disorder; Parkinson's disease;
 neuropathic pain; gene; ss.

Homo sapiens.

WO2003046167-A1.

05-JUN-2003.

26-NOV-2002; 2002WO-EP013268.

27-NOV-2001; 2001US-0333131P.

(FARB) BAYER AG.

Koehler RH;

WPI; 2003-505196/47.

XX New polynucleotide encoding a serine/threonine protein kinase
 PT polypeptide, useful for diagnosing, preventing or treating diseases
 PT associated with serine/threonine protein kinase dysfunction, e.g. cancer
 or diabetes.
 XX
 PS Disclosure; Fig 5; 196pp; English.

XX The present sequence represents a human serine/threonine protein kinase
 CC (I) related nucleotide sequence from the present invention. (I) has
 CC cytosolic, antidiabetic, neuroprotective, antiparkinsonian, nootropic,
 CC cerebroprotective, analgesic, antiinflammatory, hepatotropic,
 CC gastrointestinal, gynaecological, uropathic and dermatological
 CC activities, and can be used in gene therapy. Serine/threonine protein
 CC kinase polynucleotide and polypeptide sequences can be used in
 CC diagnosing, preventing, ameliorating or treating diseases associated with
 CC serine/threonine protein kinase dysfunction. They may also be used to
 CC identify test compounds that may act, for example, as activators or
 CC inhibitors at the enzyme's active site. The human serine/threonine
 CC protein kinase and its fragments are also useful in raising specific
 CC antibodies that can block the enzyme and effectively reduce its activity.
 CC Human serine/threonine protein kinase sequences can be used in the
 CC preparation of a medicament for modulating the activity of a serine/
 CC threonine protein kinase in a disease, e.g. cancer, diabetes, a central
 CC nervous system (CNS) disorder, a respiratory disorder (including chronic
 CC obstructive pulmonary disease), a cardiovascular disorder, a
 CC dermatological disorder, a gastrointestinal or liver disease, a
 CC haematological disorder, a musculoskeletal disorder, a reproductive
 CC disorder, or a urological disorder. CNS disorders may include Alzheimer's
 CC disease, Parkinson's disease, stroke or neuropathic pain

SQ Sequence 1534 BP; 407 A; 362 C; 365 G; 400 T; 0 U; 0 Other;

Query Match 90.6%; Score 1185.4; DB 9; Length 1534;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1186; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 122 AGCTAACAGACCTTAAAGAGCATATGTTCCATGATCTTCACTTCAACCCAGGGAGCTTC 181
DB 72 ATCTAACAGACCTTAAAGAGCATATGTTCCATGATCTTCACTTCAACCCAGGGAGCTTC 131
OY 182 AAGCTGCCCGTCCAGAAAGTCAAGAGTAAAGGCGAGGTAAAGCTGATGTTTTC 241
DB 132 AAGCTGCCCGTCCAGAAAGTCAAGAGTAAAGGCGAGGTAAAGCTGATGTTTTC 191
OY 242 AGAAGAGGATCTGAGGAGGCTTTCAAGTGAGGAAGGCTCCCTTTTGGGGAGCCT 301
DB 192 AGAAGAGGATCTGAGGAGGCTTTCAAGTGAGGAAGGCTCCCTTTTGGGGAGCCT 251
OY 302 CATCTTACTTGAACCTTGAAGAGTGGGTGAAGGCTTTATCGACAGTTTCAAGGGGA 361
DB 252 CATCTTACTTGAACCTTGAAGAGTGGGTGAAGGCTTTATCGACAGTTTCAAGGGGA 311
OY 362 TTAGAGAAATTAATGACAACTAGGCTTTAAAGTATCGATGAGATGAGAGGAAG 421
DB 312 TTAGAGAAATTAATGACAACTAGGCTTTAAAGTATCGATGAGATGAGAGGAAG 371
OY 422 GAGTCCCATTTACAGCTATCCGAGAAGCTTCTCTCGTAGGGTTGAAACATGCAATA 481
DB 372 GAGTCCCATTTACAGCTATCCGAGAAGCTTCTCTCGTAGGGTTGAAACATGCAATA 431
OY 482 TTGTCTCTGTGATGATATATCCACCAAAAGACACTGACATTCGTTTGAATACA 541
DB 432 TTGTCTCTGTGATGATATATCCACCAAAAGACACTGACATTCGTTTGAATACA 491
OY 542 TGCAACAGACCTGGCCAGTATATGCTCAGCATCCAGGAGGCTTATCTCATATG 601
DB 492 TGCAACAGACCTGGCCAGTATATGCTCAGCATCCAGGAGGCTTATCTCATATG 551
OY 602 TCAGACTTTTCAATGTTTCACTTTGCGGGGCTGCGGTACATCCACCAACCAAGCTTC 661
DB 552 TCAGACTTTTCAATGTTTCACTTTGCGGGGCTGCGGTACATCCACCAACCAAGCTTC 611

```

```

QY 662 TTCAAGGAGCCTGAACCTCAGACTTACTCACTCACTGAGAGAGCTGAACTGG 721
   |||
Db 612 TTCAAGGAGCCTGAACCTCAGACTTACTCACTGAGAGAGCTGAACTGG 671
   |||
QY 722 CTGATTTTGGTCTTGGCCGGGCGCAAGTCCATTCCTCCAGCCAGACATCTCTTCAAGAGTCG 781
   |||
Db 672 CTGATTTTGGTCTTGGCCGGGCGCAAGTCCATTCCTCCAGCCAGACATCTCTTCAAGAGTCG 731
   |||
QY 782 TGACCTCTGCTGACCGGCCCCCTGATGCTTGTCTGGAGGCACTGAATATCTCTGAGC 841
   |||
Db 732 TGACCTCTGCTGACCGGCCCCCTGATGCTTGTCTGGAGGCACTGAATATCTCTGAGC 791
   |||
QY 842 TGACATATAGGGGTGAGAGCTGATCTTTATGAATGTTCCAGGGTCAACCTTGTTC 901
   |||
Db 792 TGACATATAGGGGTGAGAGCTGATCTTTATGAATGTTCCAGGGTCAACCTTGTTC 851
   |||
QY 902 CTGGGGTTTCCACATCTTTGAACAGCTGAGAAAACTGGAGAGTCTGGAGAGTCCCTA 961
   |||
Db 852 CTGGGGTTTCCACATCTTTGAACAGCTGAGAAAACTGGAGAGTCTGGAGAGTCCCTA 911
   |||
QY 962 CAGAGATATCTTGGCCGGGAGTCTCCAGCTTACTCACTCAATCCAGAAATGTTCCAC 1021
   |||
Db 912 CAGAGATATCTTGGCCGGGAGTCTCCAGCTTACTCACTCAATCCAGAAATGTTCCAC 971
   |||
QY 1022 TGCCATCGCCTGAGAGCTTATGTTGTCTGGAACAGGCTGGGCAAGGTTCCGAAAGTCG 1081
   |||
Db 972 TGCCATCGCCTGAGAGCTTATGTTGTCTGGAACAGGCTGGGCAAGGTTCCGAAAGTCG 1031
   |||
QY 1082 AAGACCTGGCCTCCAGATGCTTAAAGGCTTTCCAGAGACCGCTCTCCGCCAGGAAG 1141
   |||
Db 1032 AAGACCTGGCCTCCAGATGCTTAAAGGCTTTCCAGAGACCGCTCTCCGCCAGGAAG 1091
   |||
QY 1142 CACTTGTATATATTTTTCAGGCGCCCTGCACTCTAGCTGTACCGCTTCTGATGAAG 1201
   |||
Db 1092 CACTTGTATATATTTTTCAGGCGCCCTGCACTCTAGCTGTACCGCTTCTGATGAAG 1151
   |||
QY 1202 AGCTTTGTTTACAGTTTTCAGAGTGAAGGCTTAAAGCAGAAATGTGTGAACCTTTGAGCCT 1261
   |||
Db 1152 AGCTTTGTTTACAGTTTTCAGAGTGAAGGCTTAAAGCAGAAATGTGTGAACCTTTGAGCCT 1211
   |||
QY 1262 CCTACGAAAGGTCACCAACCCAGCCAGTTTGAAGAAATGCTGTGA 1308
   |||
Db 1212 CCTACGAAAGGTCACCAACCCAGCCAGTTTGAAGAAATGCTGTGA 1258
   |||

```

RESULT 9

ACC79967
ID ACC79967 standard; cDNA; 1534 BP.

AC79967;
AC79967;

09-SEP-2003 (first entry)

Human serine/threonine protein kinase encoding cDNA SEQ ID NO:1.

Human; serine/threonine protein kinase; kinase; enzyme; cytosolic;
antidiabetic; neuroprotective; nootropic; antiparkinsonian; analgesic;
cerebroprotective; antiinflammatory; gastrointestinal; hepatotropic;
gynaecological; uropathic; dermatological; gene therapy; cancer;
diabetes; central nervous system disorder; CNS disorder; liver disease;
respiratory disorder; chronic obstructive pulmonary disease; stroke;
cardiovascular disorder; dermatological disorder; urological disorder;
gastrointestinal disease; haematological disorder; Alzheimer's disease;
musculoskeletal disorder; reproductive disorder; Parkinson's disease;
neuropathic pain; gene; ss.

OS Homo sapiens.

Key Location/Qualifiers

PH 104..1258

FT /tag= a

FT /product= "serine/threonine protein kinase"

XX

PN MO2003046167-A1.

XX 05-JUN-2003.

XX 26-NOV-2002; 2002W0-EP013268.

XX 27-NOV-2001; 2001US-0333131P.

XX (FARB) BAYER AG.

XX Koehler RH;

XX MPI: 2003-505196/47.

XX P-PSDB; ABR57357.

PT New polynucleotide encoding a serine/threonine protein kinase
PT polypeptide, useful for diagnosing, preventing or treating diseases
PT associated with serine/threonine protein kinase dysfunction, e.g. cancer
or diabetes.

XX Claim 1; Page 171-173; 196pp; English.

XX The present sequence encodes a human serine/threonine protein kinase (1).
XX (1) has cytosolic, antidiabetic, neuroprotective, antiparkinsonian,
XX nootropic, cerebroprotective, analgesic, antiinflammatory, hepatotropic,
XX gastrointestinal, gynaecological, uropathic and dermatological
XX activities, and can be used in gene therapy. Serine/threonine protein
XX kinase polynucleotide and polypeptide sequences can be used in
XX diagnosing, preventing, ameliorating or treating diseases associated with
XX serine/threonine protein kinase dysfunction. They may also be used to
XX identify test compounds that may act, for example, as activators or
XX inhibitors at the enzyme's active site. The human serine/threonine
XX protein kinase and its fragments are also useful in raising specific
XX antibodies that can block the enzyme and effectively reduce its activity.
XX Human serine/threonine protein kinase sequences can be used in the
XX preparation of a medicament for modulating the activity of a serine/
XX threonine protein kinase in a disease, e.g. cancer, diabetes, a central
XX nervous system (CNS) disorder, a respiratory disorder (including chronic
XX obstructive pulmonary disease), a cardiovascular disorder, a
XX dermatological disorder, a gastrointestinal or liver disease, a
XX haematological disorder, a musculoskeletal disorder, a reproductive
XX disorder, or a urological disorder. CNS disorders may include Alzheimer's
XX disease, Parkinson's disease, stroke or neuropathic pain

XX Sequence 1534 BP; 407 A; 362 C; 365 G; 400 T; 0 U; 0 Other;

Query Match 90.6%; Score 1185.4; DB 9; Length 1534;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1186; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 122 AGCTAACGACCTTAAAGAGCATATGTTTCATATCTTCACTTTTACCCGAGGACTTC 181
   |||
Db 72 ATCTAACGACCTTAAAGAGCATATGTTTCATATCTTCACTTTTACCCGAGGACTTC 131
   |||
QY 182 AAGTGGCCCGTGGCCAGAAAGTTCAAGATAAAGGCCAGGAGTAAAGTATGTTTTC 241
   |||
Db 132 AAGTGGCCCGTGGCCAGAAAGTTCAAGATAAAGGCCAGGAGTAAAGTATGTTTTC 191
   |||
QY 242 AGAAGAGAGATCTGAGGAGGCTTTCAGTGAAGAGAGGCTCCCTTTGGGAGGAGCT 301
   |||
Db 192 AGAAGAGAGATCTGAGGAGGCTTTCAGTGAAGAGAGGCTCCCTTTGGGAGGAGCT 251
   |||
QY 302 CATCTTACTTGAACCTTGAAGAGAGTGGGTGAAGGCTCTTATGCGACAGTTTCAAGGGGA 361
   |||
Db 252 CATCTTACTTGAACCTTGAAGAGAGTGGGTGAAGGCTCTTATGCGACAGTTTCAAGGGGA 311
   |||
QY 362 TTAGCAGATTAATGAGCAACTAGTGCCTTAAAGTCAAGCATGAGTCAAGAGAG 421
   |||
Db 312 TTAGCAGATTAATGAGCAACTAGTGCCTTAAAGTCAAGCATGAGTCAAGAGAG 371
   |||
QY 422 GAGTCCCATTTTACAGCTATCCGAGAGGCTTCTCTGAAAGGTTTGAACATGCAATA 481
   |||
Db 372 GAGTCCCATTTTACAGCTATCCGAGAGGCTTCTCTGAAAGGTTTGAACATGCAATA 431
   |||

```

```

QY 482 TTGTGCTCTGTCATGATATATCCACACCAAGAGACCTGACATTCGTTTTGAATACA 541
Db 432 TTGTGCTCTGTCATGATATATCCACACCAAGAGACCTGACATTCGTTTTGAATACA 491
QY 542 TGCACACAGACCTGGCCAGATATATGCTCAGACATCCAGAGGGCTTCATCCTCATATAG 601
Db 492 TGCACACAGACCTGGCCAGATATATGCTCAGACATCCAGAGGGCTTCATCCTCATATAG 551
QY 602 TCAGACTTTTCATGTTTCAACTTTTTCGCGGGCCCTGGCGTATCATCAACAACAACAGTTTC 661
Db 552 TCAGACTTTTCATGTTTCAACTTTTTCGCGGGCCCTGGCGTATCATCAACAACAACAGTTTC 611
QY 662 TTTCACAGGAGACCTGAAACCTCAGAACTTACTCATCTGCTGAGAGAGCTAAACTGG 721
Db 612 TTTCACAGGAGACCTGAAACCTCAGAACTTACTCATCTGCTGAGAGAGCTAAACTGG 671
QY 722 CTGATTTTGGTCTTGGCCGGGGCCAGTTCATCCAGCCAGACATCTCTTGAGAAAGTGG 781
Db 672 CTGATTTTGGTCTTGGCCGGGGCCAGTTCATCCAGCCAGACATCTCTTGAGAAAGTGG 731
QY 782 TGACCTCTGTCATACCGCCCTGATGCTTGTGCTGGAGGCACTGAATATTCCTCTGAGC 841
Db 732 TGACCTCTGTCATACCGCCCTGATGCTTGTGCTGGAGGCACTGAATATTCCTCTGAGC 791
QY 842 TGGACATATGGGGTGCAGGCTGCATCTTTATTTGAAATGTTTCCAGGGTCAACTTTGTTTC 901
Db 792 TGGACATATGGGGTGCAGGCTGCATCTTTATTTGAAATGTTTCCAGGGTCAACTTTGTTTC 851
QY 902 CTGGGGTTTCCAAACATCTTTGAAACAGCTGGAGAAATCTGGGGAGTCTGGGAGTCCCTA 961
Db 852 CTGGGGTTTCCAAACATCTTTGAAACAGCTGGAGAAATCTGGGGAGTCTGGGAGTCCCTA 911
QY 962 CAGAGATATCTTGGCCGGAGTCTCCAGCTACTTAACATCAATCCAGATGTTTCCAC 1021
Db 912 CAGAGATATCTTGGCCGGAGTCTCCAGCTACTTAACATCAATCCAGATGTTTCCAC 971
QY 1022 TGGCTACGCTTGGAGACCTTCATGTTGTCTGAAACAGGCTGGGAGGGTTCCTGAAAGCTG 1081
Db 972 TGGCTACGCTTGGAGACCTTCATGTTGTCTGAAACAGGCTGGGAGGGTTCCTGAAAGCTG 1031
QY 1082 AAGACCTGCGCTCCAGATGCTAAAGGCTTCCAGAGACCGCGTCTCCGCCCAAGAAAG 1141
Db 1032 AAGACCTGCGCTCCAGATGCTAAAGGCTTCCAGAGACCGCGTCTCCGCCCAAGAAAG 1091
QY 1142 CACTTGTTCATGATATATTTCAAGCGCCCTGCATCTCAGCTGTACCAAGCTTCTGATAGG 1201
Db 1092 CACTTGTTCATGATATATTTCAAGCGCCCTGCATCTCAGCTGTACCAAGCTTCTGATAGG 1151
QY 1202 AGCTTTGTTTACAGTTTCAAGAGTGAAGGCTAAAGCCAGAAATGTGTGAACCTTTTGCCCT 1261
Db 1152 AGCTTTGTTTACAGTTTCAAGAGTGAAGGCTAAAGCCAGAAATGTGTGAACCTTTTGCCCT 1211
QY 1262 CCTACAGAAAGGTCAACCAAGCCAGCTTGAAGCAAAATGCTGTGA 1308
Db 1212 CCTACAGAAAGGTCAACCAAGCCAGCTTGAAGCAAAATGCTGTGA 1258

```

```

RESULT 10
AAL48889
ID AAL48889 standard; cDNA; 2203 BP.
XX
XX AAL48889;
AC
XX
XX 24-OCT-2002 (first entry)
DT
XX
XX Human pfcaltre family kinase splice form 1 coding sequence.
DE
XX
XX Human; pfcaltre family kinase; kinase; enzyme; testis; brain; cytosolic;
KW uterus endometrium adenocarcinoma; lung fibroblast; splice form;
XX kidney renal cell adenocarcinoma; gene therapy; gene; ss.
OS Homo sapiens.

```

```

XX
FH Key Location/Qualifiers
FT 5'UTR e1.117
FT CDS /tag= a
FT 118..1149
FT /tag= b
FT /product= "kinase"
FT 3'UTR 1149..2203
FT /tag= c
XX
XX MO200261060-A2.
XX
XX 08-AUG-2002.
XX
XX 17-JAN-2002; 2002MOC-US001106.
XX
XX 31-JAN-2001; 2001US-0265151P.
XX
XX 09-MAR-2001; 2001US-00801861.
XX
XX (PEKE ) PE CORP NY.
XX
XX Yan C, Ketchum K, Di Francesco V, Beasley EM;
XX
XX WPI; 2002-608515/65.
XX
XX P-PSDB; AAO18613.
XX
XX New human kinase peptide and nucleic acid molecule, useful for treating
XX PT disorders associated with abnormal expression of kinase protein, e.g.
XX PT adenocarcinoma of uterus or lung, in drug screening assays and
XX PT pharmacogenomic analysis.
XX
XX PS Claim 4; Fig 1; 131pp; English.
XX
XX CC The present invention provides the protein, cDNA and gene sequences of
XX CC two splice variants of a human pfcaltre family kinase. The sequences are
XX CC specifically expressed in the human testis, brain, uterus endometrium
XX CC adenocarcinoma, lung fibroblasts, kidney renal cell adenocarcinoma, and
XX CC can be used to treat related diseases. The present sequence is the cDNA
XX CC of splice variant 1 of the invention
XX
XX SQ Sequence 2203 BP; 657 A; 460 C; 538 G; 548 T; 0 U; 0 Other;
XX
XX Query Match 77.1%; Score 1009; DB 6; Length 2203;
XX Best Local Similarity 100.0%; Pred. No. 2.4e-312; Indels 0; Gaps 0;
XX Matches 1009; Conservative 0; Mismatches 0;
XX
XX QY 1 ATGGGTCAAGAGCTGTGTGCAAGAAGCTGTACAGCTGATGAGAGCTGCTACCATTTGTA 60
XX Db 118 ATGGGTCAAGAGCTGTGTGCAAGAAGCTGTACAGCTGATGAGAGCTGCTACCATTTGTA 177
XX
XX QY 61 GAGGAGGAGGAGGACACAGCTGTGAGAGAGTCAAGCCAGAGAGCTGCGTTTC 120
XX Db 178 GAGGAGGAGGAGGACACAGCTGTGAGAGAGTCAAGCCAGAGAGCTGCGTTTC 237
XX
XX QY 121 AAGCTTACAGACCTTAAAGAGATCATATGTCATGACTTCAATTTTCAACCCAGGAGACTT 180
XX Db 238 AAGCTTACAGACCTTAAAGAGATCATATGTCATGACTTCAATTTTCAACCCAGGAGACTT 297
XX
XX QY 181 CAAGCTGCCCGTGGCCAGAAAGTTCAAGAGTAAAGGCGACGAGTAAAGATGTTGTTT 240
XX Db 298 CAAGCTGCCCGTGGCCAGAAAGTTCAAGAGTAAAGGCGACGAGTAAAGATGTTGTTT 357
XX
XX QY 241 CAGGAAGAGATCTGAGCGAGGGTTTTCAGTGGAGAGAGAGCTCCCTTTTGGGGAGGCC 300
XX Db 358 CAGGAAGAGATCTGAGCGAGGGTTTTCAGTGGAGAGAGAGCTCCCTTTTGGGGAGGCC 417
XX
XX QY 301 TCATCTTACTTGAACCTTGAAGAGCTGGGTGAAGGCTCTTAAGCCACATTTTCAAGGGG 360
XX Db 418 TCATCTTACTTGAACCTTGAAGAGCTGGGTGAAGGCTCTTAAGCCACATTTTCAAGGGG 477
XX
XX QY 361 ATTAGCAATTAATGACCAACTAGTGGCTTTAAAGTCAATGAGATGAATGACAGAGAA 420
XX Db 478 ATTAGCAATTAATGACCAACTAGTGGCTTTAAAGTCAATGAGATGAATGACAGAGAA 537

```

```

QY 421 GGAGTCCCATTTTACAGCTATCCGAGAGCTTCTCTCTGAGAGGTTTGAACATGCCAAT 480
Db 538 GGAGTCCCATTTTACAGCTATCCGAGAGCTTCTCTCTGAGAGGTTTGAACATGCCAAT 597
QY 481 ATGTGCTCTCGATGATGATATCCACACCAAGAGACATGATCGTTTTGAATAC 540
Db 598 ATGTGCTCTCGATGATGATATCCACACCAAGAGACATGATCGTTTTGAATAC 657
QY 541 ATGACACAGACCTGCGCCAGATATGTCCTCAGCATCAGAGGGCTTCAATCTCATTAAT 600
Db 658 ATGACACAGACCTGCGCCAGATATGTCCTCAGCATCAGAGGGCTTCAATCTCATTAAT 717
QY 601 GTGAGACTTTTCAATGTTTCAATTTTGGGGGCTGGCCGATTCACATCCACCAACAGTT 660
Db 718 GTGAGACTTTTCAATGTTTCAATTTTGGGGGCTGGCCGATTCACATCCACCAACAGTT 777
QY 661 CTTCACAGGAGCTGAAACCTCAGAACTTACTCATGATCAGCTGGAGAGCTCAAACTG 720
Db 778 CTTCACAGGAGCTGAAACCTCAGAACTTACTCATGATCAGCTGGAGAGCTCAAACTG 837
QY 721 GCTGATTTTGTCTTGGCCGAGCCCAAGTCCATTCACAGCAATATCTTCAGAAATC 780
Db 838 GCTGATTTTGTCTTGGCCGAGCCCAAGTCCATTCACAGCAATATCTTCAGAAATC 897
QY 781 GTGACCTCTGTGTAACCGGCCCCCTGATGCTTTGTGGAGACCATGAAATTTCTCTGAG 840
Db 898 GTGACCTCTGTGTAACCGGCCCCCTGATGCTTTGTGGAGACCATGAAATTTCTCTGAG 957
QY 841 CTGACATATGAGGATGAGGCTGATCTTATTGAAATGTTCCAGGGTCAACCTTTGTTT 900
Db 958 CTGACATATGAGGATGAGGCTGATCTTATTGAAATGTTCCAGGGTCAACCTTTGTTT 1017
QY 901 CTTGGGCTTTTCAACATCTTTGAAACAGCTGGAAGAAATCTGGAGGTTCTGGAGTCCCT 960
Db 1018 CTTGGGCTTTTCAACATCTTTGAAACAGCTGGAAGAAATCTGGAGGTTCTGGAGTCCCT 1077
QY 961 ACAGAGGATCTTGGCCGAGGCTCCCAAGTACCTCACTCACTCACTCACTCACTCACT 1009
Db 1078 ACAGAGGATCTTGGCCGAGGCTCCCAAGTACCTCACTCACTCACTCACTCACTCACT 1126

RESULT 11
ABA00722 standard; cDNA; 1376 BP.
AC ABA00722;
XX 04-MAR-2003 (first entry)
DT Human KPP-7 cDNA, Incyte ID No. 7494145CB1.
XX
DE
XX
XX
KW Gene, kinase; phosphatase; KPP; cell proliferation; arteriosclerosis;
KW atherosclerosis; cirrhosis; hepatitis; reproduction; infertility;
KW paroxysmal nocturnal haemoglobinuria; polycythaemia vera; psoriasis;
KW primary thrombocytopenia; cancer; development; renal tubular acidosis;
KW anaemia; mental retardation; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; epilepsy; Sjogren's syndrome; uveitis; asthma;
KW menstrual cycle; autoimmune; inflammation; rheumatoid arthritis; AIDS;
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; allergy;
KW diabetes mellitus; glomerulonephritis; Goodpasture's syndrome; gout;
KW Graves' disease; Hashimoto's thyroiditis; irritable bowel syndrome;
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
KW Reiter's syndrome; mouse; PTPAIRE kinase; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH CDS 298..1263
FT /tag= a
FT /product= "KPP-7"
XX
XX WO200290530-A2.

```

```

XX 14-NOV-2002.
PD
XX
XX 16-JAN-2002; 2002WC-US001369.
PF
XX
XX 18-JAN-2001; 2001US-0263083P.
PR 23-FEB-2001; 2001US-0271117P.
PR 23-FEB-2001; 2001US-02711205P.
PR 16-MAR-2001; 2001US-0276859P.
PR 23-MAR-2001; 2001US-0278504P.
PR 23-MAR-2001; 2001US-0278522P.
PR 29-MAR-2001; 2001US-0280266P.
PR 29-MAR-2001; 2001US-0280510P.
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX
XX Lee EA, Walla NK, Baughn MR, Ison CH, Gururajan R, Arvien C;
PI Yao MG, Jackson JL, Tang TY, Yue H, Tran B, Ding L, Lu DM;
PI Lal PG, Warren BA;
XX
XX MPI; 2003-111972/10.
DR P-PSDB; AAG79729.
DR
XX
XX The sequences given in ABA00716-23 encode human kinases and phosphatases
CC (KPP). The KPP polypeptides, polynucleotides, and agonists and
CC antagonists to them, are useful for diagnosing, treating or preventing
CC disorders associated with aberrant expression of KPP, particularly cell
CC proliferative disorders (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer's disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's
CC disease, diabetes mellitus, glomerulonephritis, Goodpasture's syndrome,
CC gout, Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome,
CC multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's
CC syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,
CC bacterial, fungal, parasitic, protozoan or helminthic infections. They
CC are also useful in the assessment of the effects of exogenous compounds
CC on the expression of nucleic acid and amino acid sequences of proteins
CC associated with KPP. The protein encoded by this sequence is homologous
CC to mouse PTPAIRE kinase
XX
XX Sequence 1376 BP; 374 A; 321 C; 342 G; 339 T; 0 U; 0 Other;
SQ
XX
XX Query Match 67.9%; Score 887.6; DB 8; Length 1376;
XX Best Local Similarity 99.6%; Pred. No. 1.7e-273;
XX Matches 890; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
QY 122 ACCTAACAGACCTTAAGAGAGCATGATTCATGATCTTATTCACCCAGGGAGCTTC 181
Db 266 ACTTAACAGACCTTAAGAGAGCATGATTCATGATCTTATTCACCCAGGGAGCTTC 325
QY 182 AGCTGCCCGTCCAGAAAGTTCAAGAGTAAAGGCCACAGAGTAACAGTATGTTTTC 241
Db 326 AAGCTGCCCGTCCAGAAAGTTCAAGAGTAAAGGCCACAGAGTAACAGTATGTTTTC 385
QY 242 AGGAAGAGATCTGAGGAGGAGGTTTTCAGTGGAGGAAGAGCTCCCTTTGGGGAGGCT 301
Db 386 AGGAAGAGATCTGAGGAGGAGGTTTTCAGTGGAGGAAGAGCTCCCTTTGGGGAGGCT 445
QY 302 CATCTTACTTGAACCTTGGAAGAGCTGGGTGAAGGCTCTTATTCGACAGTTTCAAGGGGA 361
Db 446 CATCTTACTTGAACCTTGGAAGAGCTGGGTGAAGGCTCTTATTCGACAGTTTCAAGGGGA 505

```